UNIVERSITY OF CALIFORNIA

Santa Barbara

Resonance Enhanced Multi-photon Spectroscopy of DNA

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemistry

by

Marshall Robert Ligare

Committee in charge:

Professor Mattanjah deVries, Chair

Professor Michael T. Bowers

Professor Steven K. Buratto

Professor Peter C. Ford

March 2015

The dissertation of Marshall Robert Ligare is approved.

_____

Michael T. Bowers

_____

Steven K. Buratto

_____

Peter C. Ford

_____

Mattanjah de Vries, Committee Chair

March 2015

Resonance Enhanced Multi-photon Spectroscopy of DNA

by

Marshall Robert Ligare

ACKNOWLEDGEMENTS

First and foremost I would like to thank Professor Mattanjah deVries for providing me with this opportunity and all of the knowledge that he has shared with me over these years. I would like to thank all of the people that I have worked with here at UCSB. The time that I have spent has been intellectually challenging but also lots of fun thanks to all of you. My former group members Nathan Svadlenak and Lisa Gulian thank you for all that you taught me in lab and all the hard work that you have contributed to what is presented here as well as the friendship and great conversations we had at work. All of my current group members Shawn Owens, Faady Siouri, Jake Berenbiem, Sam Boldissar and the undergraduates Paul Mazzella, Carmen, Chris and Corey thanks for all your help guys it really takes a team to keep all the equipment running and I appreciate your hard work. To Louis Grace and Alexander Mikhailovsky I owe a special thanks in helping me with all the parts of the instruments, physics, electronics and engineering that I did not have experience with when I started. Without your help I am not sure that I could have done this work so thank you very much. I would like to also thank my graduate committee members, Professor Steve Buratto, Professor Peter Ford and Professor Mike Bowers for all of your advice and mentorship over the years.

I would like to thank my family for providing me with this opportunity and for their support along the way. I would like to specially thank my wife Samantha for helping me take care of life when I was too busy with work and putting up with my crazy schedule! I could not have done this without you.

VITA OF MARSHALL ROBERT LIGARE

March 2015

Education

University of California Santa Barbara — Current PhD Candidate

- Physical / Analytical Chemistry
- GPA 3.97

Humboldt State University — Bachelor of Science

- General Chemistry
- GPA 3.39

Santa Barbara City College

- Math / General Education
- GPA 3.41

Research Experience

University of California Santa Barbara (2009-2014)

- Currently I work with Professor Mattanjah S. de Vries. We combine laser spectroscopy with mass spectrometry to study excited state dynamics and structure of gas phase biomolecules. I have studied a range of single molecules, clusters and ions using varied multiple photon techniques including: Resonance enhanced multi-photon ionization (REMPI), Infrared multi-photon dissociation (IRMPD), IR-UV dip spectroscopy, UV-UV dip spectroscopy, IR-UV gain spectroscopy.

- Application of high resolution spectroscopy/mass spectrometry to objects of antiquity. Pottery samples from both Maya and Mississippi native peoples were analyzed for tea and chocolate type beverages. Caffeine, theobromine and theophylline were identified in the samples.

- A collaboration with the FOM Institute in the Netherlands allowed me to study the hydrogen bonding competition of the nucleobase and phosphate moieties of nucleotide monophosphate cluster ions. This was done with an in house built Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS) in combination with the mid infrared free electron laser (FELIX)

Humboldt State University (2004-2008)

- Professor Matthew P. Hurst and I built an analytical lab. We set up a class A clean room for trace metal analysis as well I assembled, used and maintained a differential pulse anodic stripping voltammeter for trace metal analysis.

- Worked with Professor Scott R. J. Oliver at UC Santa Cruz (SURF 2008) where we used rare earth metals, Erbium and Gadolinium, to develop novel metal organic frameworks using the hydrothermal method.

- Worked with Professor Robert W. Zoellner investigating tautomeric energies of substituted Pyrrole. This was done using Spartan and Gaussian computational methods.

- Under the direction of Professor William F. Wood I collected and preformed organic extractions on two local plants. *Tolmiea menziesii* and *Disporum smithii*, were both shown to have banana slug antifeedant activity by the compound 2,6-Nonadienal. The compound was identified and quantified using gas chromatography/mass spectrometry.

- Developed a micro-scale radical chlorination experiment for Professor Joshua Smith's Organic Chemistry II lab.

Instrumentation / Lab Skills

Instruments: GC/MS, LC/MS, XRD, Single Crystal X-Ray, NMR, UV-Vis, FTIR, Raman, TGA/MS, TOF/MS, FTICR/MS, Electron Microscopy, Differential Anodic Stripping Voltammetry, Cyclic Voltammetry

Vacuum Equipment: Rotary vain pumps, diffusion pumps, turbo pumps and cryo-pumps

Lasers and Optics: Solid State Nd:YAG nanosecond and picosecond lasers, femtosecond Ti-Sapphire oscillator and amplifier systems, gas phase CO_2 and Excimer lasers, alignment of dye lasers, OPO/OPA, and harmonic generation systems, routinely aligned two and three laser experiments.

Programs: Mathematica, Gaussian, Spartan, Molden, AutoCAD, LabView, Origin, Excel, PowerPoint, Word.

Teaching/Work Experience

University of California Santa Barbara September 2012 - February 2013

Analytical Chemistry TA: I was responsible for helping undergraduate students with homework questions, creating exam questions and grading exams.

University of California Santa Barbara June 2011 - June 2012

Mass Spectrometry Facility Assistant: Maintained a Waters electrospray Q-TOF tandem mass spectrometer. I regularly calibrated the instrument and monitored it for performance. I helped graduate students prepare samples as well as performed accurate mass measurements for publication in organic journals.

University of California Santa Barbara September 2009 - June 2010

General Chemistry Lab Instructor: Prepared and taught general chemistry labs. I graded lab reports and helped undergraduates with general lab techniques

Humboldt State University January 2006 - December 2008

Student Assistant in Chemistry Stockroom: Prepared instruments/apparatus for general chemistry I and II labs. Prepared unknowns for qualitative analysis and stock solutions of acids and bases for experiments.

Humboldt State University August 2007 - May 2008

Supplementary Course Instructor: Prepared and gave one hour lectures for General Chemistry I and II courses

Humboldt State University January 2006 - May 2007

Drop-in Tutor for General Chemistry Courses: Helped students with homework and chemistry concepts

Awards

Phi Lambda Upsilon Award for academic achievement	2009-2010
Weiss Analytical Chemistry Scholarship Award	2007
President of the Chemistry Society at Humboldt State University	2006-2007
American Chemical Society Analytical Chemist of the Year	2005-2006

Presentations

Oral Presentation: SCOPE outreach program Ventura High School. "My Journey from High School Dropout to PhD" (2012, 2013)

Poster: UCSB Alumni Event "Shedding Light on Prebiotic Chemistry" Ligare, Marshall R.; Gulian, Lisa E.; Svadlenak, Nathan. (2013)

Oral Presentation: 1st Annual Graduate Student Symposium "Laser Mass Spectrometry: Probing the Origins of Life and Analytical Methods for Cultural Heritage Objects" (2012)

Poster: Faraday Discussion 150, Basel Switzerland. "Spectroscopic Methods for Hot Ground State Detection of Organic Molecules" Ligare, Marshall R.; Gulian, Lisa E.; Svadlenak, Nathan. (2011)

Publications

Ligare, M. R.; Siouri, F. M.; Nachtigallova, D.; de Vries, M. S.* Identification of the Long Lived Dark State of Bare Nucleobases Thymine and Uracil as Triplet Character Through IR and UV Spectroscopy. In Progress, Submitting *to Journal of the American Chemical Society*

Ligare, M. R.; Siouri, F. M.; Nachtigallova, D.; de Vries, M. S.* Excited State Dynamics of Thymine and Uracil Probed Through the Long Lived Dark States. In Progress, Submitting to *Physical Chemistry Chemical Physics*.

Ligare, M.R.; Siouri, F. M.; Mazzella, P. A.; Identification of Caffeine, Theophylline and Theobromine in North American Pottery Samples Through Direct Spectroscopic and Mass analysis. In Progress, Submitting to *Analytical Chemistry*.

Gulian, L. E.; Ligare, M. R.; Owens, S. X.; Smith S. T.*; de Vries, M.S.* Identification of Cacao use in Mayan Ceremonial Drinking Vessels. In Progress, Submitting to *Analytical Chemistry*.

Ligare, M. R.; Rijs, A. M.; Berden, G.; Kabelac, M.; Nachtigallova, D.; Oomens, J.* de Vries, M. S.* Resonant IRMPD of Nucleotide Monophosphate Anionic Clusters. In Progress, Submitting to *Photochemistry Photobiology*.

Wood, W. F.; Ligare, M. R., (2E,6Z)-2,6-Nonadienal a banana slug antifeedant from crushed leaves of Tolmiea menziesii and Disporum smithii. *Biochem. Syst. Ecol.* **2008,** *36* (11), 875-876.

Wilson, W. R.; Ligare, M.R.; Codding, S.J.; Berumen, J. M.; Zoellner, R.W.* SUBSTITUENT EFFECTS ON THE RELATIVE ELECTRONIC ENERGIES OF PYRROLE AND THE IMINE TAUTOMERS OF PYRROLE. *Journal of Undergraduate Chemistry Research.* **2007.** 6. 119

References

Mattanjah S. de Vries, Professor, University of California Santa Barbara

- Email: devries@chem.ucsb.edu
- Phone: 805-893-5921

Michael T. Bowers, Professor, University of California Santa Barbara

- Email: bowers@chem.ucsb.edu
- Phone: 805-893-2893

Peter C. Ford, Professor, University of California Santa Barbara

- Email: ford@chem.ucsb.edu
- Phone: 805-893-2443

Steven K. Buratto, Professor, University of California Santa Barbara

- Email: buratto@chem.ucsb.edu
- Phone: 805-893-3393

ABSTRACT

Resonance Enhanced Multi-photon Spectroscopy of DNA

by

Marshall Robert Ligare

For over 50 years DNA has been studied to better understand its connection to life and evolution. These past experiments have led to our understanding of its structure and function in the biological environment but the interaction of DNA with UV radiation at the molecular level is still not very well understood. Unique mechanisms in nucleobase chromaphores protect us from adverse chemical reactions after UV absorption. Studying these processes can help develop theories for prebiotic chemistry and the possibility of alternative forms of DNA. Using resonance enhanced multi-photon spectroscopic techniques in the gas phase allow for the structure and dynamics of individual nucleobases to be studied in detail. Experiments studying different levels of structure/complexity with relation to their biological function are presented.

Resonant IR multiphoton dissociation spectroscopy in conjunction with molecular mechanics and DFT calculations are used to determine gas phase structures of anionic nucleotide clusters. A comparison of the identified structures with known biological function shows how the hydrogen bonding of the nucleotides and their clusters free of solvent create favorable structures for quick incorporation into enzymes such as DNA polymerase.

Resonance enhanced multi-photon ionization (REMPI) spectroscopy techniques such as resonant two photon ionization (R2PI) and IR-UV double resonance are used to further elucidate the structure and excited state dynamics of the bare nucleobases thymine and uracil. Both exhibit long lived excited electronic states that have been implicated in DNA photolesions which can ultimately lead to melanoma and carcinoma. Our experimental data in comparison with many quantum chemical calculations suggest a new picture for the dynamics of thymine and uracil in the gas phase. A high probability of UV absorption from a vibrationally hot ground state to the excited electronic state shows that the stability of thymine and uracil comes from its intrinsic molecular properties and possibly a hydrogen bonding solvent capable of dissipating excess vibrational energy.

Due to the high specificity and sensitivity of resonant two photon ionization coupled with molecular beam mass spectrometry a new analytical technique for identifying molecular markers in archaeological vessels is presented. The xanthine alkaloids theobromine, theophylline and caffeine are identified in Central American and North American pottery sherds by direct desorption/resonant laser ionization mass spectrometry.

Table of Contents

I. Introduction: Gas Phase Spectroscopy of Biomolecules

1. Motivation

The complexity of biological systems can be overwhelming at the cellular and sub-cellular level. Even just a single protein or enzyme can be comprised of hundreds or thousands of individual amino acids that work together to make the overall structure. The four nucleotides, adenosine, guanosine, cytidine and thymidine, along with a fifth uridine, comprise the information for all known life. The diversity of life comes from many levels of structure and complexity that just four molecules provide. The current hierarchy of molecular structure from a Biochemistry standpoint is comprised of four levels. The primary structure, which is the order or sequence of a polymer such as nucleotides in single stranded DNA or amino acids in a protein. The secondary structure, which describes the hydrogen bonding and van der waals interactions, for example bonding between bases of single or double stranded DNA. Tertiary structure, which describes the larger scale three dimensional space which comes about from the combination of primary and secondary structures. And finally quaternary structure, which describe the interaction of tertiary structures, for example the biding of DNA polymerase with DNA. This structure - function relationship is of utmost importance in the biological environment. Only 20 primary amino acids and 5 nucleotides comprise over 99% all proteins, DNA and RNA. Experiments designed to study these systems are very difficult and are limited in their detail due to the complexity of the biological environment. Because of these limitations fewer experiments have been able to focus on how the individual building blocks of proteins and DNA contribute to the overall structure and function of these larger systems. Here we describe the experimental methods

and give examples of how gas phase experiments can contribute to a better understanding of the structure-function relationship that lies beneath the primary level in biological systems.

2. Objectives

1. Extend gas phase infrared spectroscopy of isolated nucleobases to 2'-deoxynucleotide-5'-monophosphate anions and anionic clusters $[dAMP\text{-}dTMP - H]^{1-}$, $[dGMP\text{-}dCMP - H]^{1-}$.

2. Identify tautomeric structures for thymine and uracil in a laser desorption seeded molecular beam. Identification of the structure and character of the long lived excited state or "dark" state for thymine and uracil.

3. Gain insight to the excited state dynamics of thymine and uracil by wavelength dependent lifetime measurements and multiple pump-probe techniques.

4. Development of a new analytical method for the identification of analytes in complex real world samples by direct laser desorption, REMPI mass spectrometry and its application to Archaeometry.

II. Resonant Infrared Multi-photon Dissociation Spectroscopy of Nucleotide Monophosphate Anions and Anionic Clusters $[dAMP\text{-}dTMP - H]^{1-}$ and $[dGMP\text{-}dCMP - H]^{1-}$

Introduction

The importance of DNA and its biological function as the genetic code of life was shown about 70 years ago. Since then, there have been countless experiments to obtain a better understanding of this biological polymer. More recently experimentalists have taken a reductionist approach to focus on the intrinsic properties of the building blocks of DNA. These experiments revealed that the nucleobases themselves can exhibit ultra-short electronic excited state lifetimes that vary with structure.[1] Experiments probing the structure of the neutral G-C and A-T nucleobase pairs in the gas phase have reported several planar hydrogen bonding motifs other than the Watson-Crick (WC) structure that is prevalent in DNA. The WC structure was observed only for the substituted 1-methylcytosine-9-ethylguanine cluster.[2] The failure to observe the WC structure for bare G-C pairs is ascribed to a short-lived excited state of the order of a picosecond, which precludes 2 photon ionization by nanosecond laser pulses. This short lifetime is the result of ultrafast internal conversion, via conical intersections and may be important in protecting the nucleobases against potentially harmful photochemistry.[3] For A-T pairs the WC structure is not necessarily the most stable one and it is also not observed in gas phase experiments with nanosecond two photon ionization.[4] Solution phase experiments aimed at studying base

pairing interactions typically use non-polar solvents and substituted nucleosides to minimize hydrogen bonding competition from the solvent and to promote hydrogen bonding rather than stacking of the base pairs.[5,6] These experiments have revealed that under those conditions many base pairing motifs are possible, for example A-T can form up to four structures that occur with almost equal probability. A similar experiment probing the excited state lifetime of the G-C base pair in chloroform has shown that the mechanism for excited state decay of the substituted nucleosides in non-polar solvent differs from that of the base pair in the gas phase.[7] These experiments connect gas phase and solution phase experiments. However, they do not include the phosphate moiety. The phosphate moiety is highly polar and capable of strong hydrogen bonds that should be able to compete with nucleobase bonding in solution as well as in the gas phase. The structures of the biologically relevant forms of DNA become increasingly difficult to identify as the systems become larger in size because more low energy structural isomers may contribute to the spectrum and because of possible increased vibrational congestion. To minimize these difficulties, vibrational mode information in the mid-IR (500 cm^{-1} – 1800 cm^{-1}) can be very useful. This region contains all hydrogen bonding modes through carbonyl, amide, amine and phosphate moieties, facilitating a complete structural analysis.

Nucleotide ions have been previously studied in the gas phase by various mass spectrometry techniques such as collision induced dissociation,[8] ion mobility,[9,10] H/D exchange[11] and IRMPD.[12,13] Gas phase studies allow for investigation of fundamental intrinsic properties of isolated biomolecular building blocks, by excluding the role of the biological environment. Electronic spectroscopy of DNA fragments in the gas phase has so-far mostly been applied to neutral nucleobases[2] and some nucleosides.[14] The combination of

resonant photoionization *via* electronic states with infrared hole burning provides opportunities for isomer specific studies. For ions, gas phase structural information has been reported for nucleobase-metal ion clusters,[15,16] protonated nucleobases[17] and nucleotide ions[13,18] using IR multiple-photon dissociation (IRMPD). Here, we extend these gas phase spectroscopic studies of DNA to the anionic DNA pairs [dGMP-d CMP - H]$^{1-}$, [dAMP-dTMP - H]$^{1-}$. We aim to gain insight in the hydrogen bonding competition between the phosphate moieties and the nucleobases by examining the vibrational frequency differences between the monomers and the clusters. We perform structural analyses by comparing with high level ab-initio calculations. The role of theory is even more critical here than in the case of neutral species studied by resonance enhanced multi-photon ionization (REMPI) spectroscopy because experimental data are not isomer specific. At the same time, when developing high level theoretical treatments, gas phase data serve as benchmarks for computational results, allowing optimization of methods and functionals.

Methods

1. Instrumentation

1.1 Electrospray Ionization

Electrospray ionization (ESI) has become the most successful and widely used technique for generating large gas phase ions for mass spectrometry. It was first conceptualized by Malcom Dole in 1968[19] and later coupled to mass spectrometry by John Fenn, who received the noble prize in 2002 for his contribution to the development of ESI and its use for mass spectrometry.[20,21] This techniques employs chemical rather than physical ionization and occurs through ionic solution being pushed through a thin metallic capillary to which an

electric field is applied. At the exit of the capillary the charged droplets are vaporized and quickly evaporate to leave behind isolated charged ions. Depending on the conditions of the ionization source; electrical potential, evaporation temperature, instrument entrance voltage ect... different ions and ionic clusters will be formed. Two major advantages to electrospray over other ionization methods are the ability two form very large ions, > 10,000 amu. The gentle nature of the ionization leads to low fragmentation, thus a range of new systems that were previously unattainable in the gas phase such as full proteins[22] and nanoparticles can be studied without solvent interactions.

1.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

A brief description regarding the theory of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) and the particular mass spectrometer used for this experiment will be given here. For a complete theoretical treatment of FT-ICR MS see.[23] The important points that will be covered here pertain to the unique advantages of FT-ICR MS for resonant IR multi-photon dissociation spectroscopy.

FT-ICR MS was first introduced in 1974 by Comisarow and Marshall.[23] The technique has since received much attention due to the very high mass resolution attainable as well as the ability to collect mass spectra without destruction of the ions. This allows for tandem mass spectrometry or MS^n, which as will be discussed below, is particularly important for the current application of mass selected IR action spectroscopy. Mass identification is accomplished by measuring the frequency at which ions cycle in a constant, applied magnetic field. By having knowledge of the magnetic field strength and measuring the frequency of ions cycling in the field, the mass to charge ratio can then be calculated by equation 1.

$$\nu_c = \frac{\omega_c}{2\pi} = \frac{(1.535611\times10^7)B_0}{m/z} \qquad 1$$

where the frequency ν is in Hz, applied magnetic field B_o is in Tesla, mass m is in atomic units and charge z is elementary charge. Typical magnetic field strengths are between 4 and

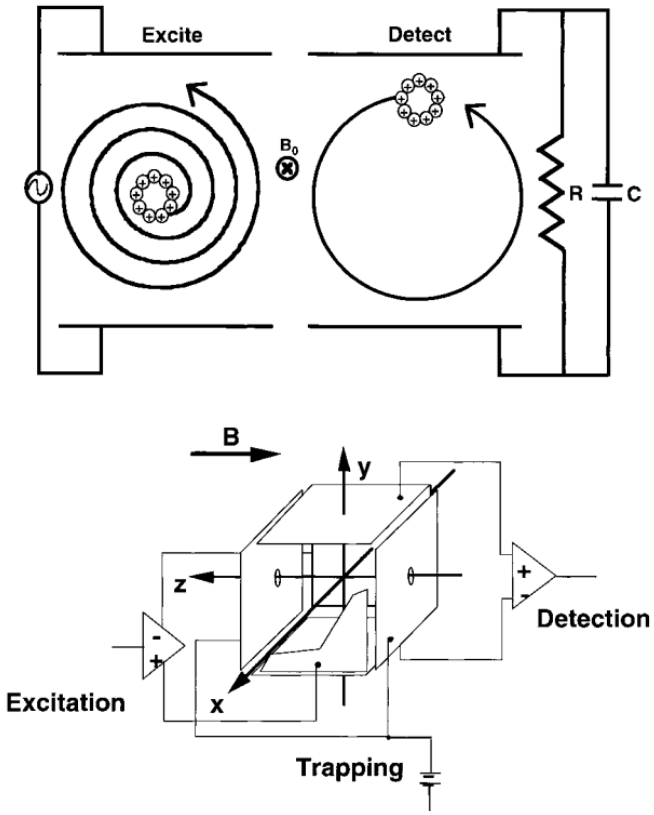


Figure 1. (Top) ICR orbital motion and excitation. (Bottom) ICR cell for trapping, excitation and detection of ion packets. Ref 23

9 Tesla which give ion cyclotron frequencies in the kHz - MHz range for mass to charge ratios of 15 -10,000. This large dynamic mass range shows another advantage of FT-ICR MS. The ICR frequency itself is not sufficient to detect a particular mass. This is due to the ions in the detection cell being non-coherent in their cyclotron motion. A separate excitation

step must be used to create a coherent packet of ions that can be detected. Figure 1 shows the excitation of an ion packet (top) and the ICR cell (bottom) that is used for trapping, excitation and detection of ions. Excitation of the ion packet can be done in many ways. An excitation frequency which is resonant with the ion cyclotron frequency of a particular mass will send the ion packet into an excited orbit around the magnetic field as shown in the top of figure 1. This creates a coherent ion packet that is cycling near the detection plates. The change in current on each plate, upper and lower, as the ions move by the plates can now be detected. An important feature of this technique is shown in equation 2.

$$r = \frac{(V_{p-p})T_{excite}}{2dB_0} \qquad 2$$

where $V_{p\text{-}p}$ is the peak to peak voltage applied to the excitation plates, T is the time period for excitation and d is the distance between the excitation plates. The excited cyclotron radius is independent of the m/z thus broadband excitation schemes like the ones shown in figure 2d and 2e can excite and detect ion packets over a range of m/z simultaneously and the signals for each m/z are linearly proportional to the ion abundance in the cell. For action spectroscopy such as IRMPD this is critically important. The spectrum is a function of the total fragmentation of the parent ion at particular wavelengths. This leads to an IR spectrum that is relatively linear with absorbance thus being somewhat comparable to a direct absorbance IR spectrum. Although one should always be aware that action spectroscopy is not a direct absorbance measurement and although the mass spectrometer has a large linear range the fragmentation of an ion at different frequencies may not be linear with absorbance.

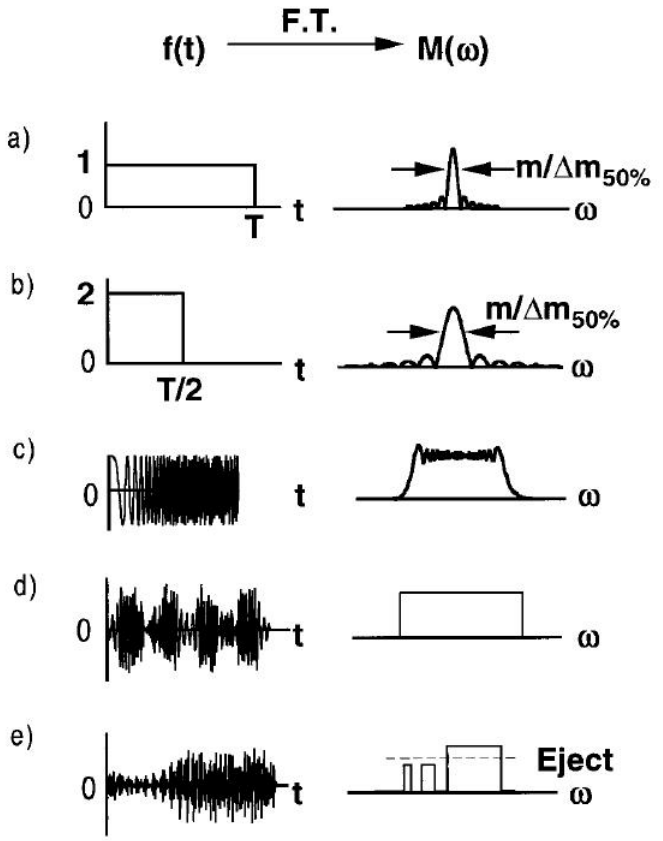


Figure 2. Time domain (left) Frequency domain (right) for ion excitation waveforms (a) single frequency excitation with time T (b) single frequency excitation with time T/2 (c) frequency chirp or "sweep" (d) SWIFT waveform (e) mass selection waveform. Ref 23.

Many different ionization sources are available which can be used in conjunction with FT-ICR MS. Both internal field and external field ionization techniques can be implemented with these instruments. Typically electron ionization, photoionization, or field ionization are used as internal field sources. While these techniques are highly useful external field techniques such as electrospray ionization (ESI) and matrix assisted laser desorption-ionization (MALDI) can introduce much larger masses with less fragmentation. These techniques introduce more molecules into the vacuum chamber and differential pumping of

the mass spectrometer is required. This is why they are generated externally of the detection region of the ICR. As well other ion optics are needed to guild the ions generated external of the magnetic field into the detection cell. For certain applications this is advantageous for mass selection steps and collisional cooling of the ions prior to injection. The instrument used for this experiment employs electrospray ionization due to the wide range of molecules that can be studied.

2. Resonant Infrared Multi-photon Dissociation Spectroscopy

This technique is quite specialized due to very large photon densities required to dissociate covalent bonds with resonant IR photons. The key here is resonant IR photons, many experimentalist use CO_2 lasers to study IR dissociation and where available quasi-resonant vibrational modes but a Free Electron Laser is the only means to cover the mid-Infrared range with enough energy to dissociate covalent bonds resonantly. So these experiments are limited by the number of facilities available with both the proper instrumentation and free electron laser. The FOM Institutes Free Electron Laser for Infrared Experiments (FELIX) in the Netherlands is just such a facility. The undulator type free electron laser creates tunable radiation in the range of 90 - 2000 cm^{-1}. The laser pulse is a 5 - 10 μs long macropulse that is a series of micropulses, 3 - 6 ps, each separated by 1 ns. The average energy of each micropulse is on the order of 10 μJ.

3. Experimental

The infrared action spectra of the anionic nucleotides, 2'-dexoynucleotide-5'-monophosphates and clusters $[dGMP\text{-}dCMP - H]^{1-}$, $[dAMP\text{-}dTMP - H]^{1-}$ are obtained by employing resonant infrared multiple-photon dissociation (IRMPD) using mid-IR photons

from the free electron laser FELIX (Free Electron Laser for Infrared Experiments).[24] The ions are trapped in a 4.7 Tesla actively shielded Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR MS) to monitor fragmentation as a function of IR frequency. The experimental setup has previously been described in detail.[25,26] A schematic of the instrument is reprinted here for clarity (Figure 3). The deprotonated nucleotide anions and anionic clusters were generated by electrospray ionization using a Waters Micromass Z-spray electrospray source which facilitates gentle ionization and vaporization of labile biomolecular species. Solutions of the mononucleotides were 1mM in 50:50 H_2O:MeOH. An equal molar addition of NaOH was added to the solution of dGMP and [dGMP-dCMP - H]$^{1-}$ to facilitate deprotonation due to low ion signal. The continuously generated ions are accumulated in a hexapole ion trap, after which they are sent through a quadrupole bender before entering an octopole ion guide that drives the ions into the FTICR cell. The anion or anionic cluster of interest is isolated using a SWIFT excitation pulse similar to the one shown in Figure 2e.[27]

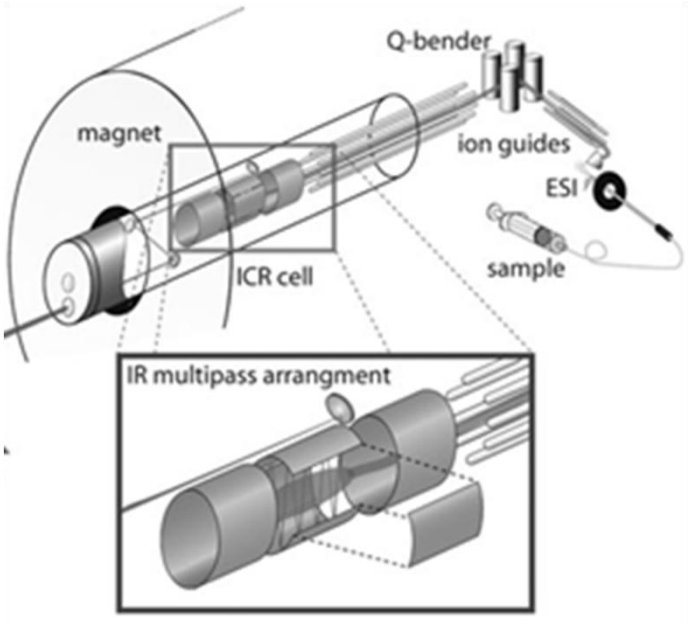


Figure 3. Schematic of the FT-ICR MS. Ref 26.

4. Computational

To explore the potential energy surfaces we used a molecular dynamics quenching method, which combines molecular dynamics with a minimization procedure of selected geometries, obtained from the trajectories. We used an on-the-fly molecular dynamics technique, employing a semi-empirical AM1 method as an external potential.[28] For the nucleotide pairs we included all possible combinations of charged and neutral nucleotides. Molecular dynamics quenching simulation was performed at 1200 K, which is sufficient to allow crossing over all relevant energy barriers and thus sample the complete potential energy surface. To avoid dissociation, we imposed harmonic restraints on the distances of centers of masses of the molecules in the complexes. The ensemble was sampled by Andersen thermostat and a total simulation time of 1 ns was used. Every picosecond the structure from the MD trajectory was minimized, with the AM1 method, and stored. The optimized structures were sorted with respect to their conformations and energies. Subsequently, the fifty most stable structures, based on relative Gibbs energies, were further optimized with the B97D[29] density functional and TZVP[30] basis set, employing the density fitting procedure. The harmonic vibrational analyses of the optimized structures were performed and the resulting frequencies without further scaling were compared with the experimental spectra. The DFT calculations employed the Gaussian 09 package.[31]

Results and Discussion

1. Monophosphate Nucleotide Anions

The IRMPD spectra of the anionic nucleotides have been previously reported by Nei et al.[13] All ground state structures and assignments here are in agreement with these earlier results. The current calculations and experimental data has been included here for three reasons: Interpretation of vibrational frequencies for the monomers is important in the assignment of the cluster structures, the calculations presented here use a different DFT functional for both the ground state energy and frequency analysis, as well we collected data at lower IR laser power yielding more resolved spectra of the high oscillator strength modes such as the carbonyl and phosphate stretches. Figures 4-7 present the IRMD spectra together with the computed spectra of the six or seven lowest energy conformations for $[dAMP - H]^{1-}$, $[dTMP - H]^{1-}$, $[dCMP - H]^{1-}$ and $[dGMP - H]^{1-}$ respectively. The structures are ordered sequentially in increasing energy. In the text structures will be referred to by number which is also sequential in increasing energy. The mid IR (500-1800 cm^{-1}) region of the vibrational action spectrum is very rich with structural information. It includes the resonances of the carbonyl stretches (around 1700 cm^{-1}, indicated in yellow), amine bending modes (around 1500 cm^{-1} indicated in green), phosphate modes (around 1250 cm^{-1}, 1050 cm^{-1}, 800-700 cm^{-1} indicted in blue) and C5'-O' stretch (around 1100 cm^{-1} indicated in red). For each $dNMP^{-}$, the five lowest energy structures calculated by DFT (Table 1) and compared with the experimental frequencies of the IRMPD spectra. As already concluded by Nei et al., for all the nucleotides the lowest energy calculated structure has the 3'OH of the ribose hydrogen bonded to the phosphate. All nucleotides except $[dGMP - H]^{1-}$ are in the anti-configurations with respect to the nucleobase and the ribose.[10,12] In all cases the

sugars line up in the C3’endo conformation with the base and phosphate lying above the plane of the sugar.

Table 1. Relative energies (ΔE, in kcal/mol) and Gibbs free energies (ΔG, in kcal/mol) for the six most stable structures of the monophosphate nucleotide anions calculated using B97D/TZVP method. The structures are sorted according to their stabilities at the free energy surface

	Adenine		Guanine		Cytosine		Thymine	
	ΔE	ΔG	ΔE	ΔG	ΔE	ΔG	ΔE	ΔG
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	-0.51	0.31	2.90	3.00	1.03	0.76	0.84	0.67
3	-0.60	0.50	4.77	3.56	3.28	3.34	1.83	1.55
4	0.94	0.66	6.19	4.84	3.84	3.43	2.94	2.30
5	2.01	1.03	5.96	6.51	5.20	3.80	2.83	3.08

[dAMP – H]$^{1-}$

The four lowest energy structures for [dAMP – H]$^{1-}$, are 39%, 23%, 16% and 13% abundant at room temperature equilibrium population. All four computed vibrational spectra form a reasonable fit to the experimental data, as shown in figure 4 The most prominent differences between the calculated spectra are in the P-OH stretch at 719 cm^{-1}. For the structure of the global minimum this mode is predicted to be the third most intense after the NH_2 scissor and the O-P-O asymmetric phosphate stretch. This intensity prediction is consistent with the experiment. All computed phosphate modes are, however, red-shifted compared to the experiments by about 50 cm^{-1}, complicating the assignment of structure 1. Structures 2, 3, 4, 5 and 6 may also be present and contribute to the experimental spectrum. All structures exhibit very good agreement between experiment and theory for the frequency of the unbound NH_2 scissor mode which is calculated without scaling ~10 cm^{-1} to the red of the experimentally observed frequency.

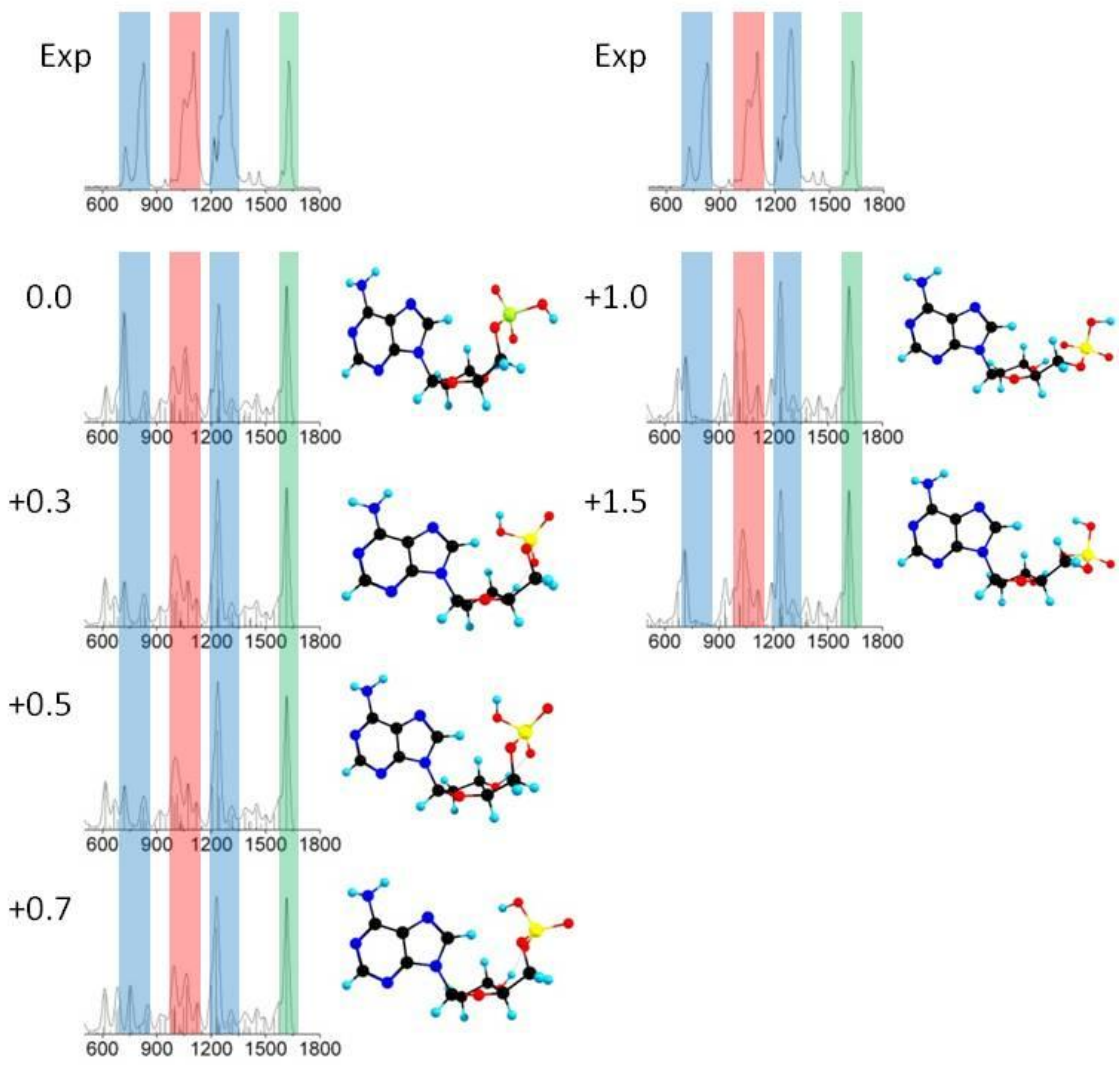


Figure 4. The six lowest energy spectra for $[dAMP - H]^{1-}$ compared with the experiment. Energies listed are Gibbs free energies in kcal/mol relative to the ground state structure

[dTMP – H]$^{1-}$

The predicted structures of 2'-deoxythymidine-5'-monophosphate are similar to those of $[dAMP - H]^{1-}$ with three nearly identical lowest energy structures, differing only in the

location and orientation of the proton bound to the phosphate. Figure 8 shows the calculated room temperature equilibrium population distribution. The two lowest energy structures have very similar vibrational spectra. Both are so similar to the experimental spectrum that excluding either structure is not possible. In fact, broadening and multiple peaks for the phosphate asymmetric stretch (1285 cm^{-1}, 1261 cm^{-1}) and P-O-H bend (1093 cm^{-1}, 1049 cm^{-1}) suggest that more than one low energy structure is present. The two calculated carbonyl stretches are predicted to be close enough in energy that they may not be resolved at the experimental bandwidth of 30 cm^{-1}. Only one peak appears in the experiment at 1709 cm^{-1}. The convoluted peak from the two carbonyl computed frequencies is centered at 1692 cm^{-1} for the global minimum and is within 2 cm^{-1} for the other low energy structures (Figure 5).

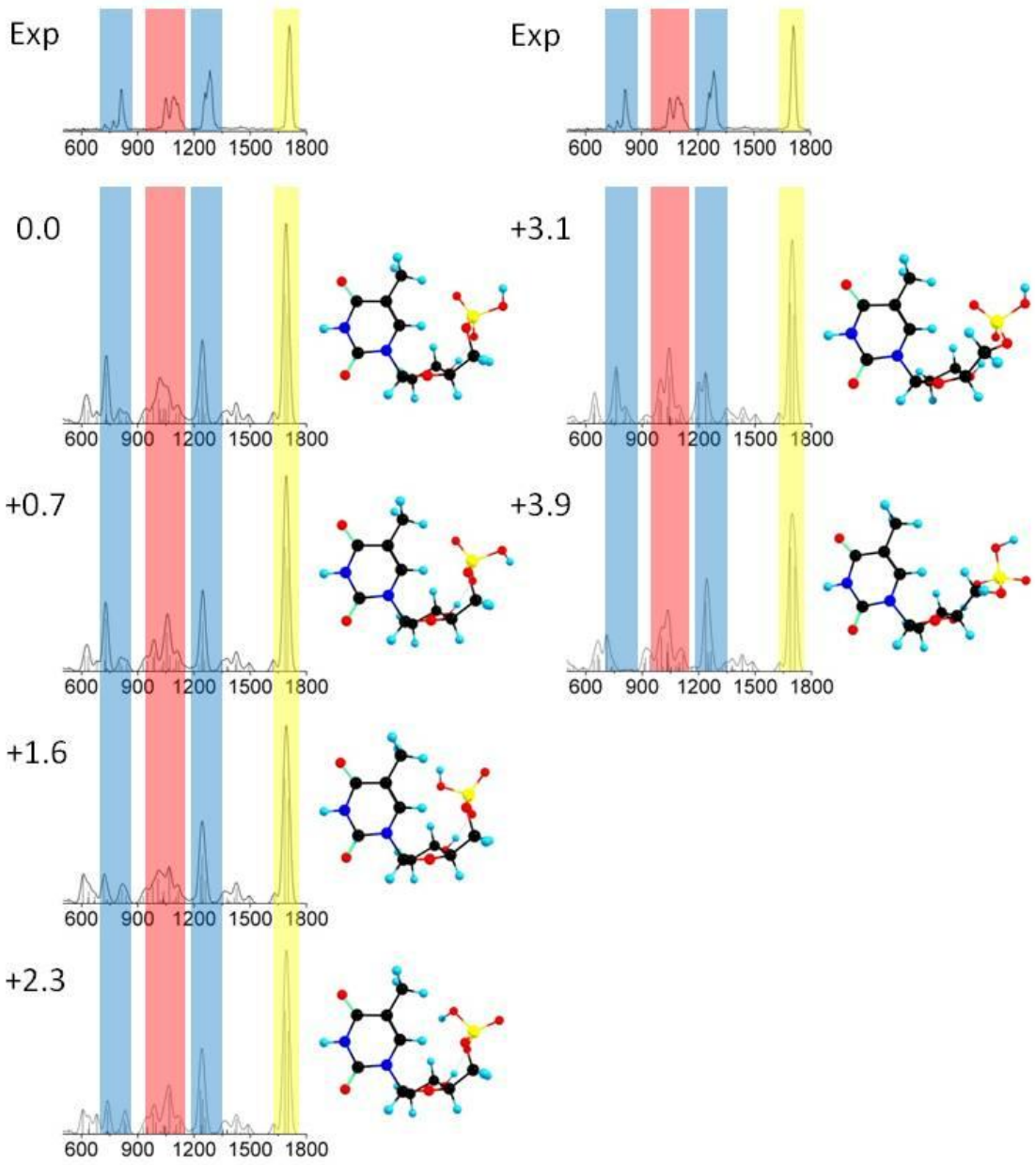


Figure 5. $[dTMP - H]^{1-}$ Experimental spectrum compared with the six lowest energy calculated spectra. Relative energies are ΔG in kcal/mol

[dCMP – H]$^{1-}$

The computed structures for 2'-deoxycytidine-5'-monophosphate contain the same set as seen for [dAMP – H]$^{1-}$ and [dTMP – H]$^{1-}$. The lowest two energy structures only differ by the orientation of the proton bonded to the phosphate. The room temperature equilibrium distribution is predicted to be 78% and 22% abundant respectively for the two lowest energy structures. There are only minor structural differences in the third lowest energy structure, which is 3.34 kcal/mol higher in energy than the global structure. The eight member ring formed by the phosphate hydrogen bonded to the 3'OH changes from a boat type structure at low energy, #'s 1,2, to a partial twist conformation at higher energies, #'s 3,4,5,6. For all three energetically lowest structures, the calculated NH_2 scissor and carbonyl modes are within ± 1 cm^{-1} of the experimental data. The spectrum for structure 7 gives us a high level of certainty that the base and sugar are in the syn configuration for all the conformers in the experiment. It also shows the upper limit of ~4.0 kJ/mol that is energetically accessible in this experiment. The NH_2 interacting with the phosphate in this structure blue shifts and the carbonyl red shifts such that the two modes are now one peak by the 30 cm^{-1} convoluted spectrum thus being the only computational spectrum that does not match the experiment in this region (Figure 6).

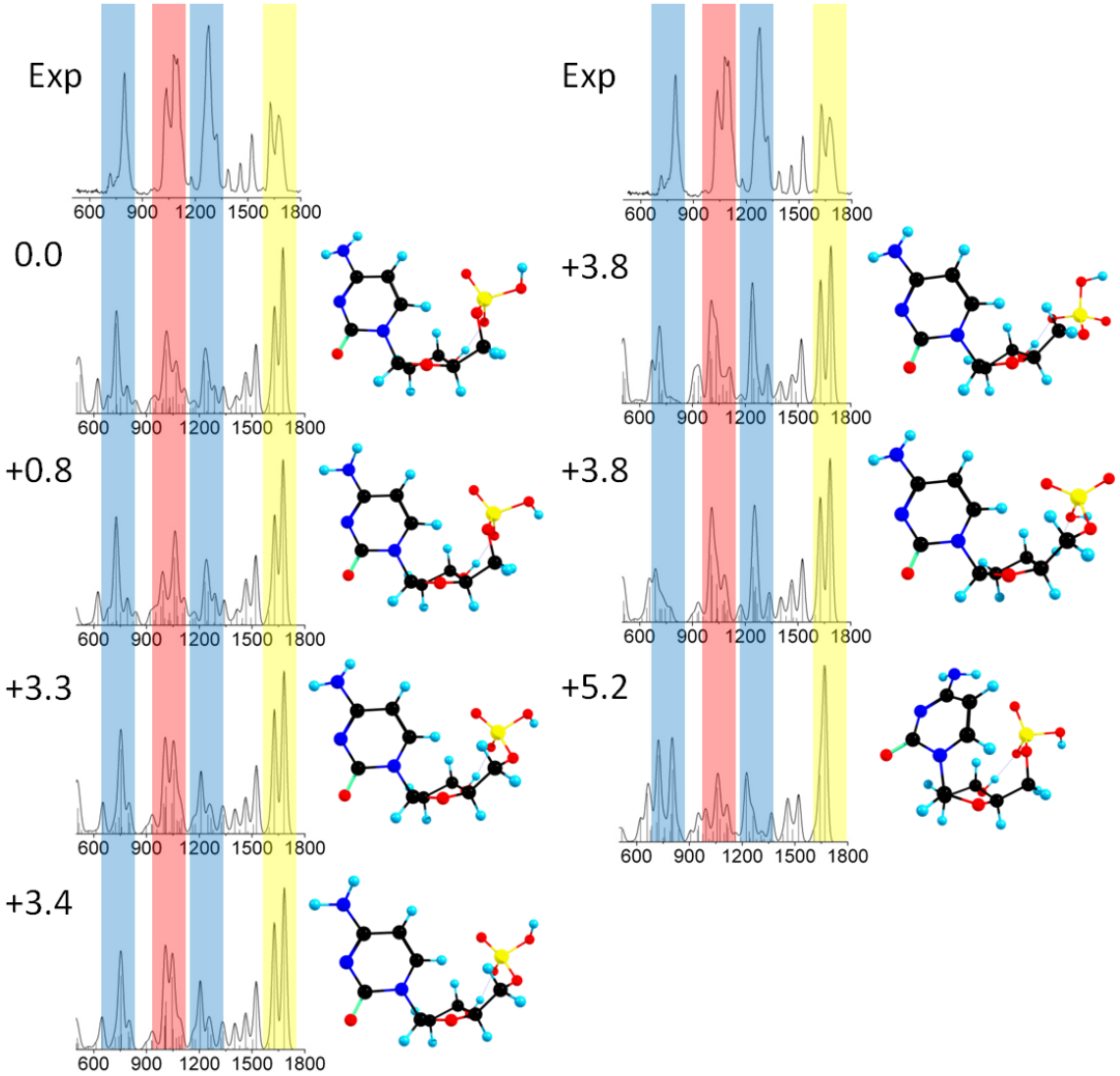


Figure 6. The seven lowest energy spectra for $[dCMP - H]^{1-}$ compared with the experiment. Energies listed are Gibbs free energies in kcal/mol relative to the ground state structure

[dGMP – H]$^{1-}$

[dGMP – H]$^{1-}$ is the only structural outlier of the four nucleotides. Its lowest energy structure is so much more stable than the energetically next higher one, that it may be expected to be predominantly populated, almost to the exclusion of all others. The second lowest energy structure is 3.0 kcal/mol higher in energy and the phosphate is still hydrogen bonded to the 3'OH of the ribose but with the protonated oxygen instead of the charged oxygen. This difference in structure changes the location and intensity of the calculated O=P-O asymmetric stretch from 1194 cm^{-1} to 1235 cm^{-1} for the doubly H-bonded structure *vs* the singly H-bonded structure respectively (Figure 7). Although the O=P-O asymmetric stretch for the higher energy structure is closer to the experimental value, the stretch at 1194 cm^{-1} from the lowest energy structure is consistent with the difference between the calculated and experimental O=P-O asymmetric modes for all the other nucleotides which is around 50 cm^{-1}. Furthermore, the relative intensity of this peak is closer to the experimental spectrum. Since both structures are in the *anti*-configuration and the NH_2 is hydrogen bonded to the phosphate, the calculated spectra are very similar; both are in good agreement with the experiment. It should be noted that the DFT functional, B97D, predicts this strongly hydrogen bound NH_2 scissor mode quite well without scaling.

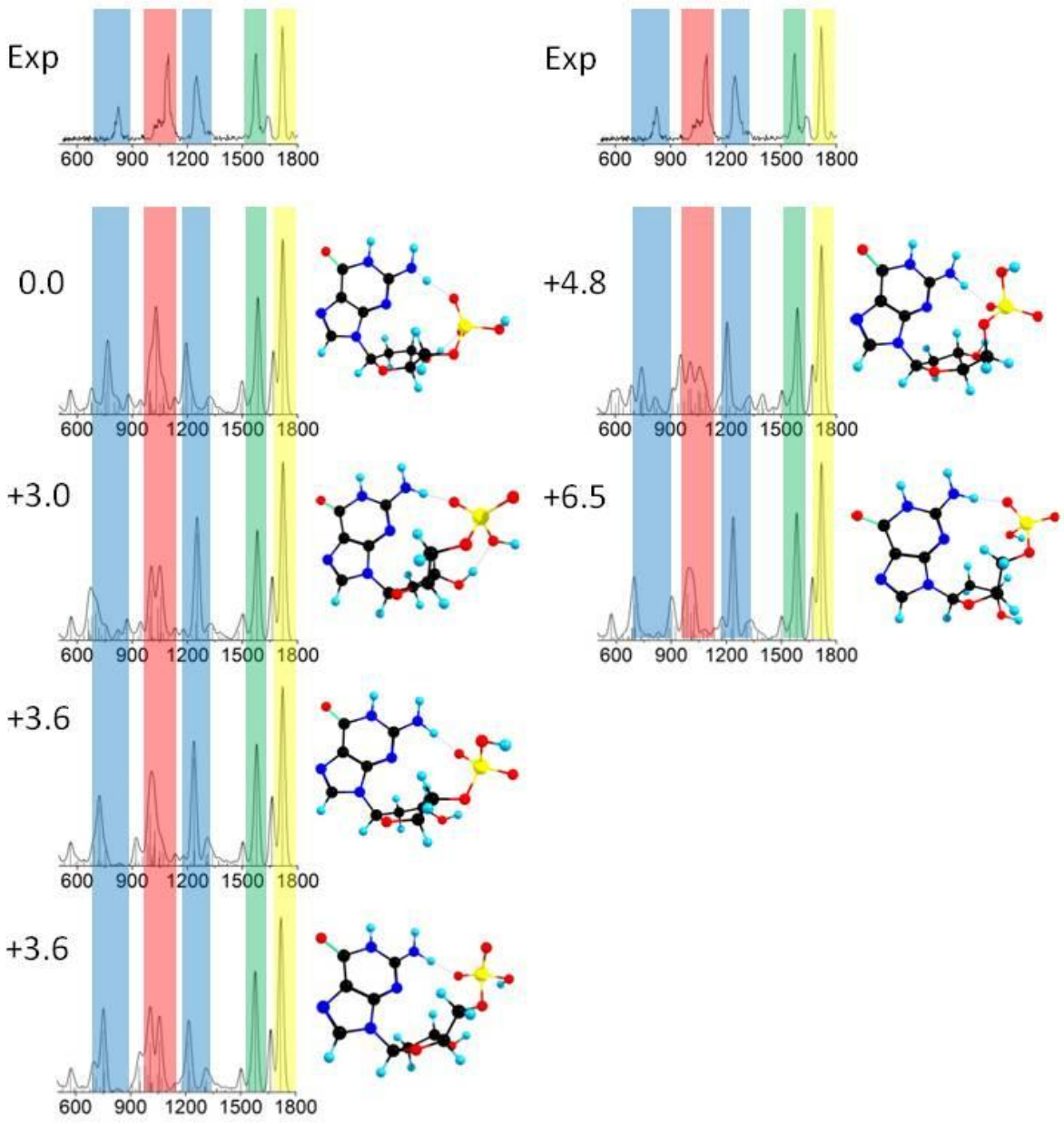


Figure 7. $[dGMP - H]^{1-}$ Experimental spectrum compared with the six lowest energy calculated spectra. Relative energies are ΔG in kcal/mol.

3.2 Anionic Monophosphate Nucleotide Clusters

In each cluster one monophosphate nucleotide is neutral and the other is negatively charged. We obtained twelve low energy calculated structures for each cluster, six where the phosphates on the pyrimidine nucleotides are negatively charged and six where the phosphates on the purine nucleotides are negatively charged. We also considered zwitterion structures for the neutral nucleotides in the clusters. dGMP has a low energy structure where the N3 position is protonated and the phosphate is deprotonated. Protonation of the bases dCMP, dAMP and dTMP were all found to be too high in energy to be considered.

All of the calculated structures show multiple hydrogen bonds, most of which occur between the phosphate moieties (Figs. 8 and 9) The structures calculated here are similar to some of the structures identified in a gas phase ion mobility study of covalently bound dinucleotides.[9] In nearly all of the structures, the two phosphate moieties are connected by at least one and in the lowest energy cases two H-bonds. Only one structure is predicted where the phosphates are H-bonded but the bases are away from each other (see Figure 8 structure 7). Our experimental data suggest that this type of structure, if present at all, is in low abundance (see below). The phosphate group can both accept and donate hydrogen bonds with its highly polar P-OH and P=O bonds. The purine and pyrimidine bases can also act as both hydrogen bond donors and acceptors. In all cases except one, the highly polar phosphates bond to each other, which aligns all clusters in the same direction, with different rotations of the base relative to the sugar. Although assignment of one single structure is not possible due to spectral congestion and the occurrence of many low energy structures, assignment of a few families of structures is possible.

[dAMP-dTMP – H]$^{1-}$

The computational data suggest that the cluster of neutral dAMP bonded to negatively charged dTMP is more stable than the cluster with the negative charge residing on dAMP. The six lowest energy structures of [dAMP-dTMP - H]$^{1-}$ fall within a range of 2.1 kcal/mol, all with the charge on the phosphate of dTMP. The lowest energy structure with the charge on dAMP is found at 4.75 kcal/mol. The structures are numbered sequentially in increasing energy.

Figure 8 shows the experimental spectrum with the calculated spectra for comparison. The B97D functional employed predicts the region above 1500 cm^{-1} quite well. Structures 1 and 2 have double hydrogen bonds from dAMP to [dTMP – H]$^{1-}$ between the phosphate groups forming an eight member ring. For all calculated structures, both ribose moieties are in the 3' endo conformation with the exception of dAMP in structure 5. Although the experimental spectrum is congested, the clusters are predicted to have intensities for the phosphate, carbonyl and amine modes that are similar to the monomers. The carbonyl vibrations on [dTMP – H]$^{-1}$, NH_2 scissor on dAMP and the various phosphate modes are color coded in the figures and are assigned in Appendix I Table 5. The two lowest energy structures fit the experimental spectrum best considering these modes. The carbonyl modes of the [dTMP – H]$^{-1}$ monomer are predicted to be only 15 cm^{-1} apart, however they are split by about twice as much in the cluster. The experimental spectrum shows some broadening of the unresolved carbonyl peaks as well as a blue shift of these vibrations compared to the monomer spectrum (Appendix 1 section 2). This shift is consistent with the calculated frequencies of structures 1, 2, 4, 5 and 6 of [dAMP-dTMP – H]$^{1-}$, where the planes of the bases are closer to perpendicular rather than parallel. In structure 3, the only stacked type

conformation with the base planes oriented parallel to each other, the symmetric and anti-symmetric carbonyl stretch modes are predicted to be much more intense and only 5 cm^{-1} apart giving a narrower carbonyl peak than seen experimentally. The NH_2 scissor mode is more intense in the experimental spectrum than predicted in any of the computed spectra, but structures 1 and 2 provide the closest match to the observed intensity and frequency of these modes.

The experiment shows that the dissociation yield is much higher for the non-covalently bound clusters than for the monomers. Therefore the spectral region from 1000-1300 cm^{-1} was collected at half the laser power of the rest of the spectrum to avoid saturation. The phosphate stretches have larger transition dipole moments and thus lower laser power is required over this range to fragment the clusters. Furthermore, the clusters are held together by H-bonds rather than by covalent bonds so a smaller number of IR photons is required to reach fragmentation for the clusters than for the monomers. The modes predicted in this region with high intensity are the O-P-O asymmetric and P=O stretches on dTMP and dAMP, respectively. The asymmetric stretch is predicted to be at 1181 cm^{-1} and the P=O stretch at 1168 cm^{-1} in structure 1. These phosphate stretches appear to the blue of the calculated frequencies analogous to the monophosphates in this experiment. Scuderi and co-workers have also reported phosphate stretches and bends to the blue of the calculated frequencies.[12] We tentatively assign the experimental peak centered at 1245 cm^{-1} to these two modes. The most intense peak in the spectrum corresponds to both C5'-OPO_3 stretches separated by 27 cm^{-1} at 991 cm^{-1} and 1018 cm^{-1}. Experimentally these peaks are at 1072 cm^{-1} and 1110 cm^{-1}, respectively. These peaks are unresolved and quite broad at the base suggesting that other modes are also contributing. Structures 3 and 4 are relatively low in

energy but do not show the same relative intensities for these sugar and phosphate modes. Therefore it is difficult to determine whether or not they contribute to the spectrum.

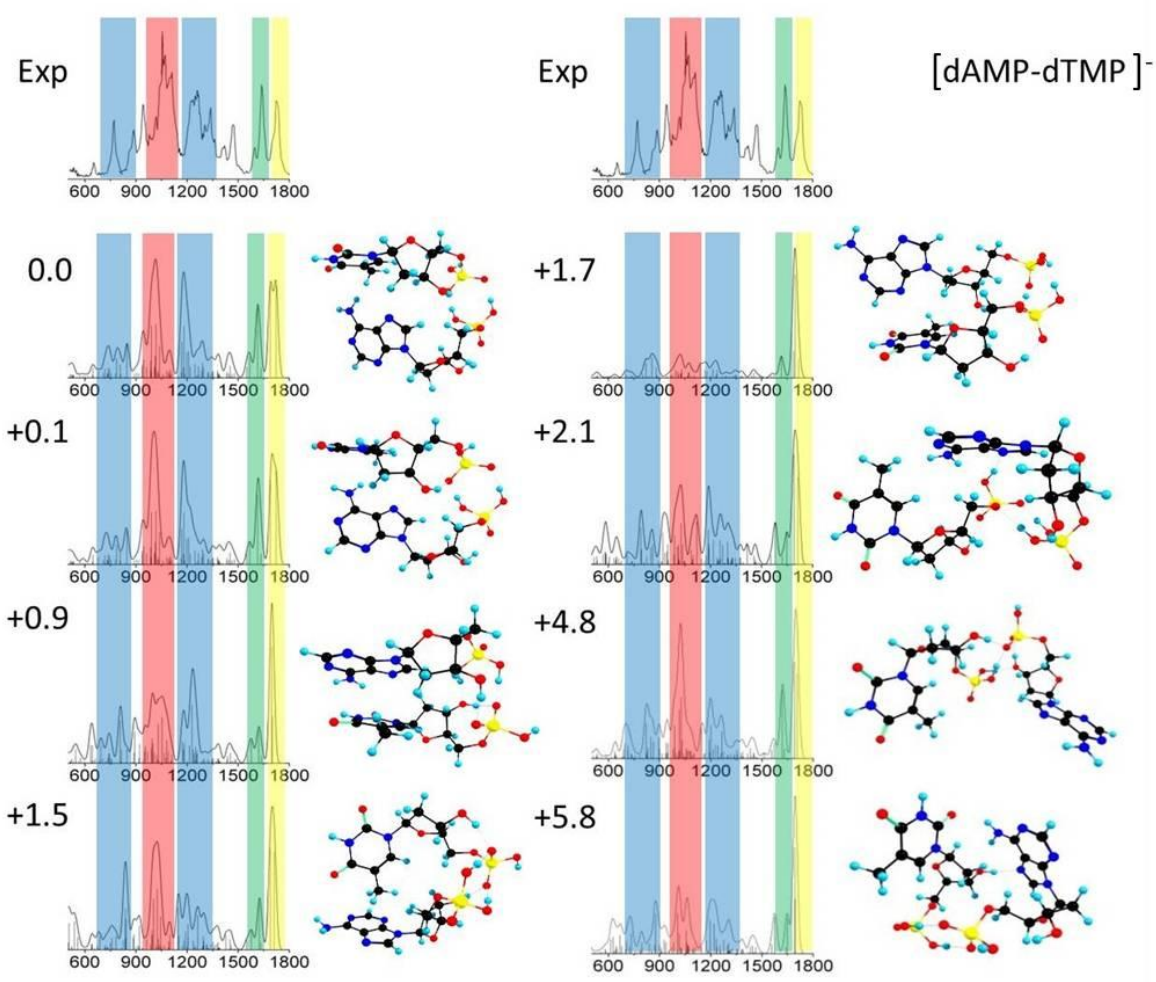


Figure 8. The eight lowest energy spectra for $[dAMP\text{-}dTMP - H]^{1-}$ compared with the experiment. Energies listed are Gibbs free energies in kcal/mol relative to the ground state structure.

[dGMP-dCMP – H]$^{1-}$

Similar to structures 1-5 of $[dAMP\text{-}dTMP - H]^{1-}$, structure 1 of $[dGMP\text{-}dCMP - H]^{1-}$ is charged on the pyrimidine nucleotide. As shown in Figure 9, the experimentally observed

spectral structure of the C=O and NH_2 modes of $[dGMP\text{-}dCMP - H]^{1-}$ is reproduced quite well by structure 1. The peak at 1703 cm^{-1} matches the C=O stretch at 1705 cm^{-1} resulting from dGMP hydrogen bonded to dCMP. In the monomer the carbonyl is unbound and appears at 1717 cm^{-1} while predicted at 1719 cm^{-1}. For the cluster, the calculation predicts that the carbonyl of dGMP is H-bonded to the NH_2 of $[dCMP - H]^-$ and therefore this mode shifts by 12 cm^{-1} to the red. This is in close agreement with the 13 cm^{-1} shift observed experimentally, from 1717 cm^{-1} to 1704 cm^{-1}. The second lowest energy structure for $[dGMP\text{-}dCMP - H]^{1-}$ lies 3.31 kcal/mol higher in energy and is predicted to be only about 0.3% abundant. Here, the carbonyl on guanine is also H-bonded to the NH_2 on dCMP and red shifted to 1699 cm^{-1}. In structure 1, the carbonyl of $[dCMP - H]^{1-}$ is H-bonded to NH_2 of guanine producing a red shift of about 15 cm^{-1} with respect to the unbound carbonyl frequency. The experimental spectrum has the largest peak in the amide region at 1644 cm^{-1}. If this is the carbonyl stretch from $[dCMP - H]^1$, it would correspond to a 30 cm^{-1} shift to the red from the unbound carbonyl frequency. In structure 2, this carbonyl is unbound and not shifted. This spectral feature most clearly distinguishes the two lowest energy structures in the region above 1600 cm^{-1} and suggests that the lowest energy structure, 1 of $[dGMP\text{-}dCMP - H]^{1-}$, is the largest contributor to the experimental spectrum. Overall the phosphate modes are calculated to the red, as was the case for $[dAMP\text{-}dTMP - H]^{1-}$, but considering these modes, the $[dGMP\text{-}dCMP - H]^{1-}$ #1 calculated structure is the one most similar to the experiment. The other, higher energy, structures shown in Figure 9 do not match the experimental spectrum in either the phosphate or carbonyl/amine regions.

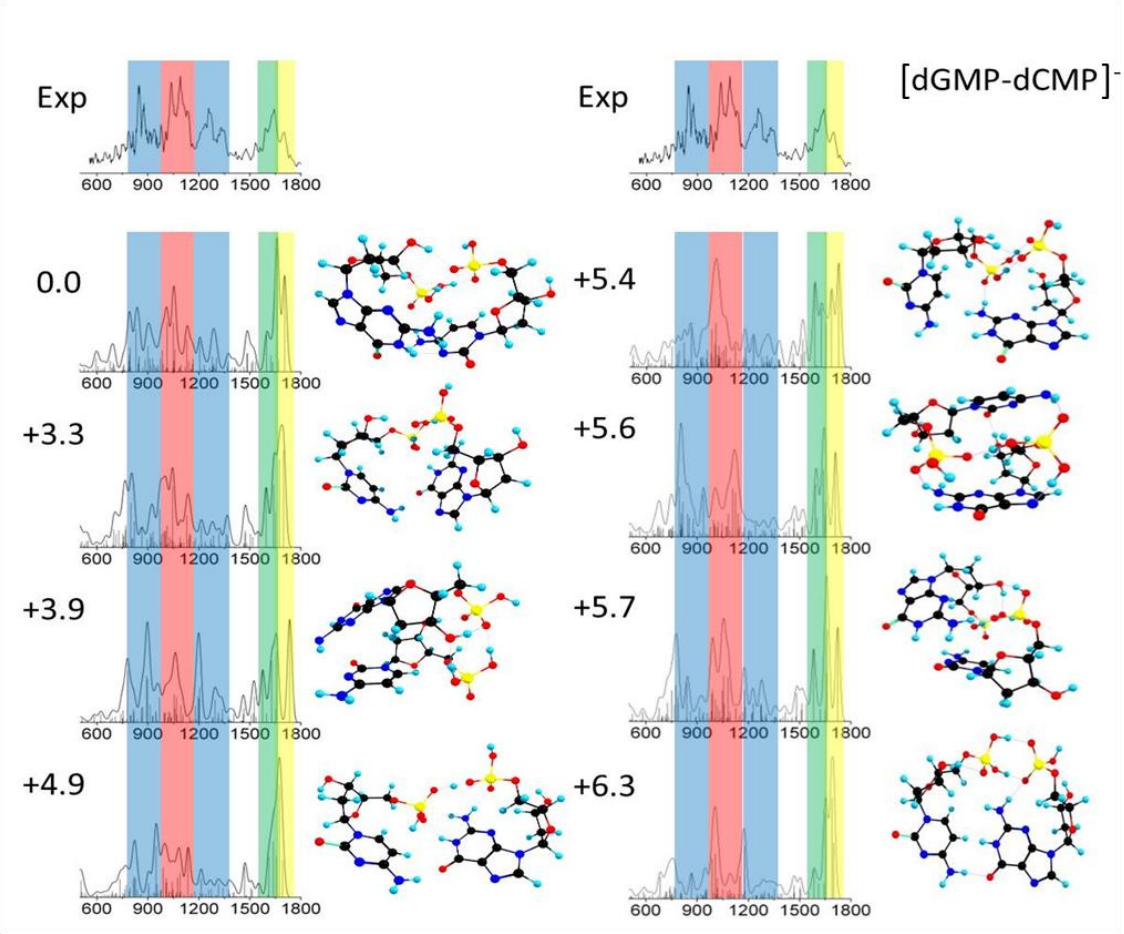


Figure 9. The eight lowest energy structures and their spectra for $[dGMP\text{-}dCMP - H]^{1-}$ compared with the experiment. The energies listed are Gibbs free energies in kcal/mol relative to the ground state structure**.**

[dGMP-dCMP – H]$^{1-}$ Cluster with Zwitterionic dGMP

The calculations also predict two low energy zwitterionic structures. In that case the N3 position on guanine is protonated and the phosphate is deprotonated leaving an overall neutral molecule that can bind to the negatively charged [dCMP – H]$^{1-}$ giving a negatively charged cluster. The spectrum for these structures is shown in Figure 10 along with the experimental spectrum.

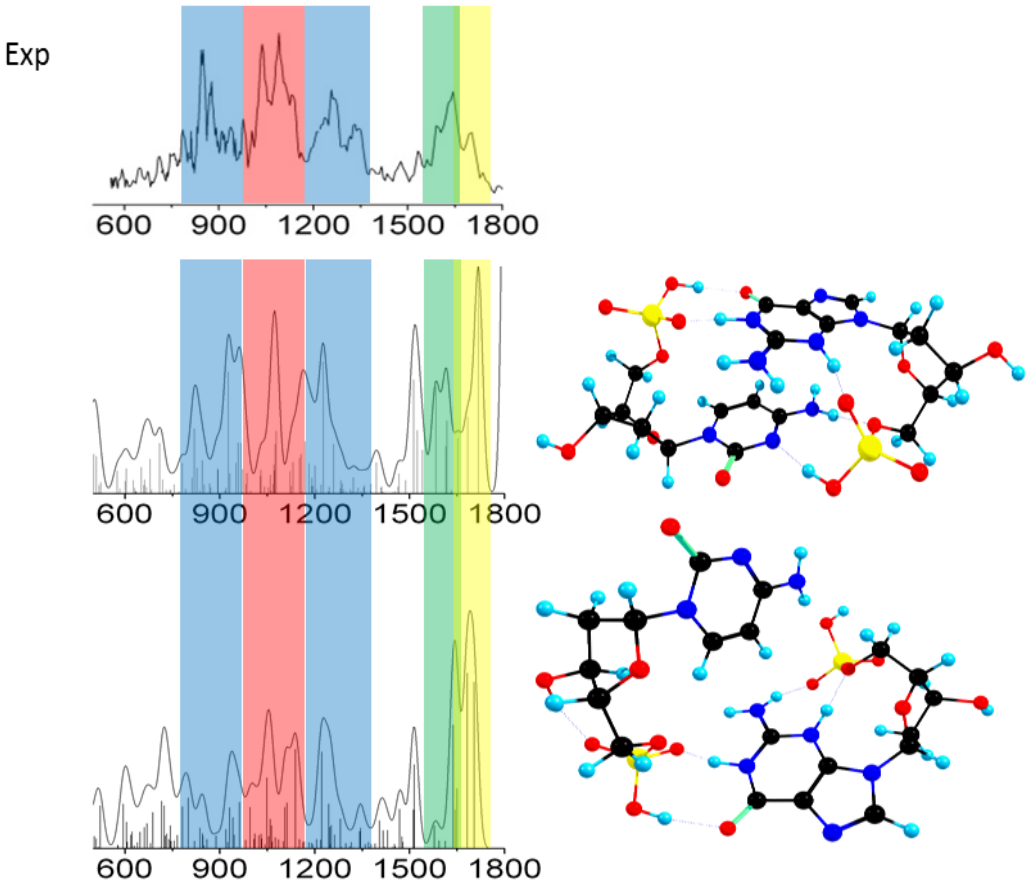


Figure 10. Comparison of experimental and calculated spectrum for zwitterionic dGMP-dCMP cluster.

A dominating feature of the spectrum comes from the N-H bends in the 1700-1720 cm^{-1} range which involves an N3H and N1H coupled symmetric in-plane bend and both are H-bonded to either their own phosphate or the other nucleotides phosphate. Furthermore, the N1H stretch is red shifted into the 1800 cm^{-1} and is predicted to have an intensity nearly three times that of any other mode below 2000 cm^{-1}. Since the experimental spectrum is limited to 1800 cm^{-1} it is difficult to tell if these structures are present. If the prediction of the intensity and location of this absorbance is within 20 cm^{-1} or so then we would expect to see a large increase in fragmentation at the 1800 cm^{-1} edge of the experimental spectrum, which is not seen. All phosphate modes are predicted to the blue of the experimental range while all other organic modes are predicted accurately or slightly to the red if hydrogen bonded. This would suggest that these strongly hydrogen bonded organic modes should be just inside of the experimental range but the overall experimental spectrum does not seem to exhibit any of the major modes predicted for these structures. Two new modes are predicted in the region between 1500-1600 cm^{-1} from a coupled in-plane asymmetric bend of the two N-H bonds. This produces the large change in the calculated spectrum for this region which also does not match the experiment.

3.3 Comparison of Gas Phase Structures with Other Experiments

IRMPD using Free-electron lasers and ICR-MS have been used in the past to obtain the IR action spectra for cyclic 3',5'-adenosine monophosphate as well as all the 2'-deoxynucleotide-5'-monophosphates.[12,13] In both these experiments the overall families of structures assigned to the spectra are the same as in this work. As mentioned in the results the phosphate modes are calculated to be lower in energy than seen experimentally while the carbonyl and amine modes are quite accurate. This was reported in the other nucleotide

experiments and with diethyl phosphate.[32] Although in the other experiments the calculation for the phosphate modes were not as red shifted but calculated frequencies for the amine and carbonyl modes were not as accurate as calculated here. Since both the nucleotide monophosphates and diethyl phosphate vibrational frequencies have previously been determined by IRMPD and calculated at the B3LYP/6-311G+(d,p) level of theory, another calculation of diethyl phosphate using B97D/TZVP level of theory is presented to compare the two functionals and basis sets on the phosphate stretches. Ethyl phosphate harmonic frequencies with both methods is given in Appendix I. As can be seen from Figure 11 the B97D/TZVP calculation predicts all modes to be lower in energy than B3LYP/6-311G+(d,p) but the phosphates are shifted even farther than the others. In the case of the major phosphate modes at ~1200 cm^{-1}, ~1000 cm^{-1}, ~700 cm^{-1}, ~600 cm^{-1} the B97D method predicts these modes 80-100 cm^{-1} to the red from the B3LYP functional. This validates the results seen for the nucleotide monophosphates for the B97D functional as it estimates the carbonyl and amine moieties accurately without scaling but underestimates the phosphates by 50-100 cm^{-1}. The major differences in the spectrum of the clusters *vs* the monomers were in the carbonyl/amine modes, thus the B97D functional is advantageous for the spectral analysis above 1400 cm^{-1} despite misrepresenting the highly polar phosphate moiety.

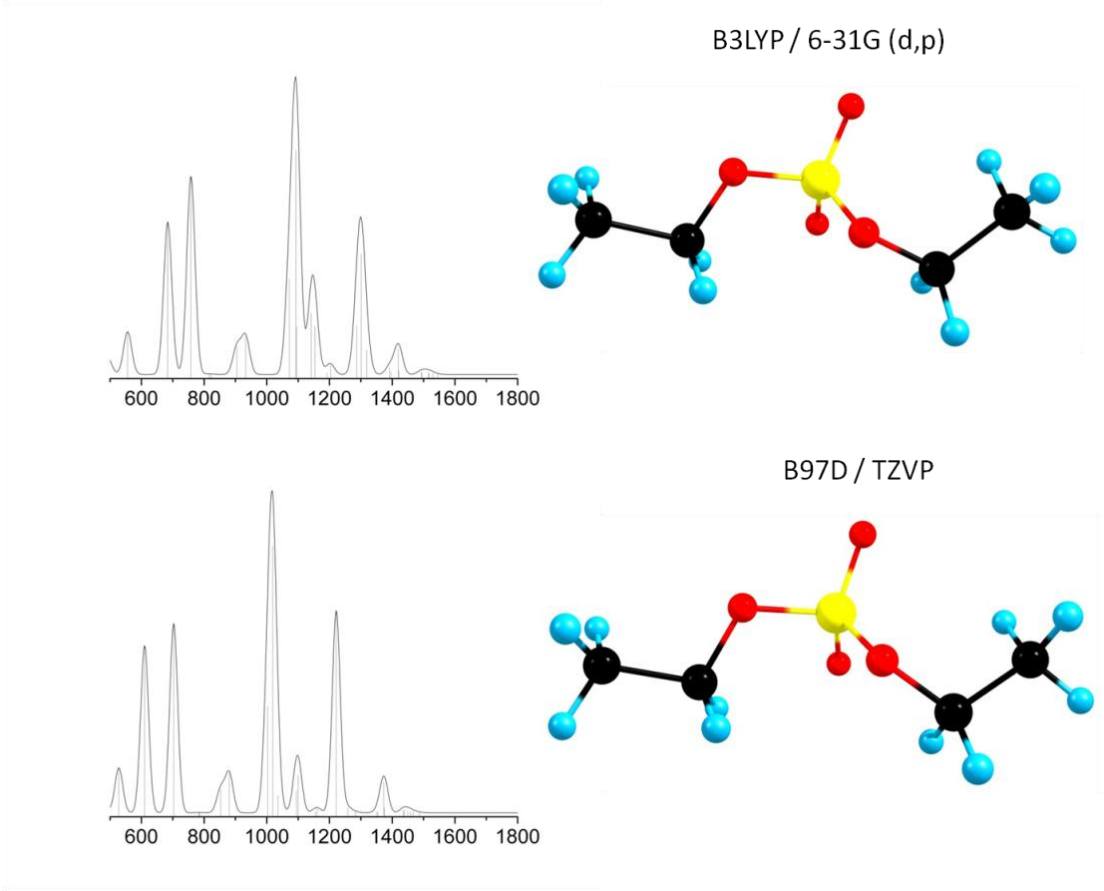


Figure 11. Harmonic frequency calculations for diethylphosphate using B3LYP/6-31G* and B97D/TZVP.

Tesmer and co-wokers have reported X-ray crystallography experiments and calculations on dATP and its binding to Adenylyl Cyclase.[33] In the binding pocket of the enzyme, the O3'H and the phosphate are bonded to the active site metal cation and this structure, pertaining to the proximity of the 3'OH and the α-phosphate, is similar to the one identified in this gas phase study on the monomers. In many of the cluster structures this preference to stabilize the phosphate charge is preserved. Other clusters show similar stabilization, however, as the base planes become parallel or "stacked" the 3'OH of one

nucleotide is hydrogen bonded to the phosphate of the other nucleotide. Thus showing a structure similar to what would be expected for DNA polymerase.

These molecules are the single units that make up the DNA polymer. DNA polymerase has one of the highest fidelity rates in biopolymerization. For fast polymerization the complex must be able to take in the monomer quickly, make small conformational changes, and add the monomer to the chain. The stacked type cluster structures predicted in this study are very similar to structures calculated for the dNTP (deoxynucleotide triphosphates) in the binding sites of DNA polymerase[34,35] as well as in the X-ray crystal structures from permanently bonded or trapped analogs to dNTPs.[36] In the active site of DNA polymerase the 3'OH end of the primer strand covalently bonds to the incoming nucleotide's α-phosphate. In all of the mononucleotides and in most of the cluster structures the 3'OH of each nucleotide is hydrogen bonded to its own phosphate to stabilize the charge. This suggests that the stabilization energy gained from this interaction is quite large yet the one cluster structure for each $[dAMP\text{-}dTMP - H]^{1-}$ and $[dGMP\text{-}dCMP - H]^{1-}$ where the bases are "stacked" the phosphates are H-bonded to each other but the 3'OH of each is now hydrogen bonded to the phosphate on the adjacent nucleotide (Figure 12). These calculations suggest that when the bases are stacked, as they are in single stranded DNA, the phosphate and 3'OH of adjacent nucleotides become proximal. This finding suggests that in the gas phase, free of solvent effects, the hydrogen bonding of the phosphates and favorable interaction of the bases directs the alignment of the nucleotides analogously to a structure optimized for quick incorporation into DNA polymerase in the biological environment.

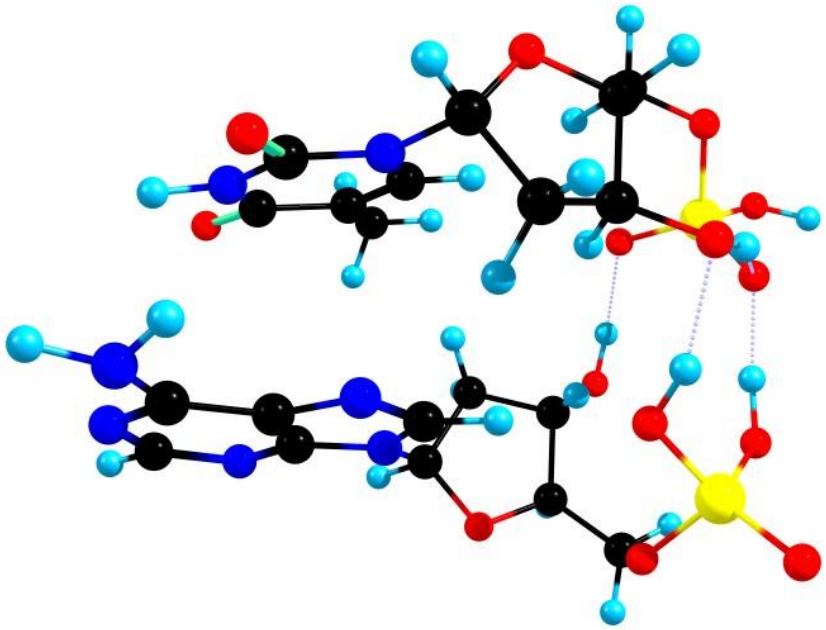


Figure 12. Structure [dAMP-dTMP – H]$^{1-}$ #3. Stacked type structure showing the preference of 3'OH for the adjacent nucleotide.

Summary

The DFT calculations predict the monomer deoxynucleotide vibrational frequencies quite well. Although the phosphate symmetric and asymmetric stretches are ~ 50 cm^{-1} to the blue of those calculated and ~ 100 cm^{-1} to the blue for the P-OH stretches, the assignments for these modes were made by their large absorbance and by comparison with other IR experiments probing phosphate[37,38] and organophosphates.[32,39] Since the carbonyl and amine frequencies for nearly all monomer structures match the experimental spectra, the structurally most diagnostic frequencies come from the phosphate. For dAMP$^-$ and dGMP$^-$ we can tentatively assign the calculated lowest energy structures to be dominant but for dCMP$^-$ the four lowest energy structures are so similar and close in energy that they all may

be present. For $dTMP^-$ the two lowest energy structures appear to be present in the experiment.

For the clusters the phosphates are always bonding with each other up to 5.5 kcal/mol and the bases may occur as either parallel or perpendicular. For $[dGMP\text{-}dCMP - H]^{1-}$ one structure at 5.6 kcal/mol occurs where the phosphates bond to the nucleobases and not to each other but this is the only example of this motif. This observation is consistent with the prediction that (1) this cluster is capable of more hydrogen bonds and (2) the $dGMP^-$ monomer is in the *syn*-conformation rather than the *anti*-conformation, giving the nucleobase the ability to be low in energy as a curled rather than extended structure. This propensity for the *syn* rather than *anti* conformation seems to be mostly conserved in the clusters and consequently a larger diversity of structures appears in the calculations of $[dGMP\text{-}dCMP - H]^{1-}$, albeit with a much higher energy gap from the lowest energy structure of $[dAMP\text{-}dTMP - H]^{1-}$. This gap allows for the higher energy structures to be more easily ruled out, while at the same time the other structures exhibit large differences in their calculated spectra.

The $[dGMP\text{-}dCMP - H]^{1-}$ cluster can form three hydrogen bonds between the nucleobases as opposed to $[dAMP\text{-}dTMP - H]^{1-}$ which can form only two hydrogen bonds. This extra H-bonding competes favorably with the bonding between the phosphate groups and thus changes the lowest energy structures from non-bonding bases in $[dAMP\text{-}dTMP - H]^{1-}$ to the stacked motifs in $[dGMP\text{-}dCMP - H]^{1-}$. A major difference appears between $[dGMP\text{-}dCMP - H]^{1-}$ and $[dAMP\text{-}dTMP - H]^{1-}$ for the preference of the charge location. While $[dAMP\text{-}dTMP - H]^{1-}$ has the six lowest energy structures with the charge on dTMP, the lowest energy for $[dGMP\text{-}dCMP - H]^{1-}$ has its charge on dCMP and for the next two low

energy structures the charge is on dGMP. There is a much larger energy gap between structure 1 and 2 giving a predicted abundance of above 99% for the lowest energy structure $[dGMP\text{-}dCMP - H]^{1-}$ #1.

References

1. Kleinermanns, K.; Nachtigallova, D.; de Vries, M. S., Excited state dynamics of DNA bases. *International Reviews in Physical Chemistry* **2013,** *32* (2), 308-342.
2. Abo-Riziq, A.; Grace, L.; Nir, E.; Kabelac, M.; Hobza, P.; de Vries, M. S., Photochemical selectivity in guanine-cytosine base-pair structures. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (1), 20-23.
3. Sobolewski, A. L.; Domcke, W.; Hattig, C., Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: The role of electron-driven proton-transfer processes. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (50), 17903-17906.
4. Plutzer, C.; Hunig, I.; Kleinermanns, K.; Nir, E.; de Vries, M. S., Pairing of isolated nucleobases: Double resonance laser spectroscopy of adenine-thymine. *ChemPhysChem* **2003,** *4* (8), 838-842.
5. Yang, M.; Szyc, L.; Rottger, K.; Fidder, H.; Nibbering, E. T. J.; Elsaesser, T.; Temps, F., Dynamics and Couplings of N-H Stretching Excitations of Guanosine-Cytidine Base Pairs in Solution. *J. Phys. Chem. B* **2011,** *115* (18), 5484-5492.
6. Greve, C.; Preketes, N. K.; Fidder, H.; Costard, R.; Koeppe, B.; Heisler, I. A.; Mukamel, S.; Temps, F.; Nibbering, E. T. J.; Elsaesser, T., N-H Stretching Excitations in Adenosine-Thymidine Base Pairs in Solution: Pair Geometries, Infrared Line Shapes, and Ultrafast Vibrational Dynamics. *J. Phys. Chem. A* **2013,** *117* (3), 594-606.
7. Biemann, L.; Kovalenko, S. A.; Kleinermanns, K.; Mahrwald, R.; Markert, M.; Improta, R., Excited State Proton Transfer Is Not Involved in the Ultrafast Deactivation of Guanine-Cytosine Pair in Solution. *J. Am. Chem. Soc.* **2011,** *133* (49), 19664-19667.
8. Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L., LOW-ENERGY COLLISION-INDUCED DISSOCIATION OF DEPROTONATED DINUCLEOTIDES - DETERMINATION OF THE ENERGETICALLY FAVORED DISSOCIATION PATHWAYS AND THE RELATIVE ACIDITIES OF THE NUCLEIC-ACID BASES. *Int. J. Mass Spectrom. Ion Process.* **1994,** *137*, 121-149.
9. Gidden, J.; Bowers, M. T., Gas-phase conformational and energetic properties of deprotonated dinucleotides. *Eur. Phys. J. D* **2002,** *20* (3), 409-419.
10. Gidden, J.; Bowers, M. T., Gas-phase conformations of deprotonated and protonated mononucleotides determined by ion mobility and theoretical modeling. *J. Phys. Chem. B* **2003,** *107* (46), 12829-12837.
11. Freitas, M. A.; Marshall, A. G., Gas phase RNA and DNA ions 2. Conformational dependence of the gas-phase H/D exchange of nucleotide-5 '-monophosphates. *J. Am. Soc. Mass Spectrom.* **2001,** *12* (7), 780-785.
12. Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Lanucara, F.; Lemaire, J.; Maitre, P.; Scuderib, D., Infrared spectroscopy of isolated nucleotides 1. The cyclic 3 ',5 '-adenosine monophosphate anion. *Int. J. Mass Spectrom.* **2008,** *270* (3), 111-117.
13. Nei, Y. W.; Hallowita, N.; Steill, J. D.; Oomens, J.; Rodgers, M. T., Infrared Multiple Photon Dissociation Action Spectroscopy of Deprotonated DNA Mononucleotides: Gas-Phase Conformations and Energetics. *J. Phys. Chem. A* **2013,** *117* (6), 1319-1335.

14. Abo-Riziq, A.; Crews, B. O.; Compagnon, I.; Oomens, J.; Meijer, G.; Von Helden, G.; Kabelac, M.; Hobza, P.; de Vries, M. S., The Mid-IR spectra of 9-ethyl guanine, guanosine, and 2-Deoxyguanosine. *J. Phys. Chem. A* **2007,** *111* (31), 7529-7536.
15. Nei, Y. W.; Akinyemi, T. E.; Kaczan, C. M.; Steill, J. D.; Berden, G.; Oomens, J.; Rodgers, M. T., Infrared multiple photon dissociation action spectroscopy of sodiated uracil and thiouracils: Effects of thioketo-substitution on gas-phase conformation. *Int. J. Mass Spectrom.* **2011,** *308* (2-3), 191-202.
16. Salpin, J. Y.; Gamiette, L.; Tortajada, J.; Besson, T.; Maitre, P., Structure of Pb2+/dCMP and Pb2+/CMP complexes as characterized by tandem mass spectrometry and IRMPD spectroscopy. *Int. J. Mass Spectrom.* **2011,** *304* (2-3), 154-164.
17. Nei, Y. W.; Akinyemi, T. E.; Steill, J. D.; Oomens, J.; Rodgers, M. T., Infrared multiple photon dissociation action spectroscopy of protonated uracil and thiouracils: Effects of thioketo-substitution on gas-phase conformation. *Int. J. Mass spectrom.* **2010,** *297* (1-3), 139-151.
18. Schinle, F.; Crider, P. E.; Vonderach, M.; Weis, P.; Hampe, O.; Kappes, M. M., Spectroscopic and theoretical investigations of adenosine 5 '-diphosphate and adenosine 5 '-triphosphate dianions in the gas phase. *Phys. Chem. Chem. Phys.* **2013,** *15* (18), 6640-6650.
19. (a) Dole, M.; Mack, L. L.; Hines, R. L., MOLECULAR BEAMS OF MACROIONS. *J. Chem. Phys.* **1968,** *49* (5), 2240-&; (b) Dole, M.; Mack, L. L.; Kralik, P., MOLECULAR BEAMS OF MACROIONS .2. *Bulletin of the American Physical Society* **1969,** *14* (3), 383-&; (c) Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M., MOLECULAR BEAMS OF MACROIONS .2. *J. Chem. Phys.* **1970,** *52* (10), 4977-&; (d) Clegg, G. A.; Dole, M., MOLECULAR BEAMS OF MACROIONS .3. ZEIN AND POLYVINYPYRROLIDONE. *Biopolymers* **1971,** *10* (5), 821-&.
20. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., ELECTROSPRAY IONIZATION FOR MASS-SPECTROMETRY OF LARGE BIOMOLECULES. *Science* **1989,** *246* (4926), 64-71.
21. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., ELECTROSPRAY IONIZATION-PRINCIPLES AND PRACTICE. *Mass Spectrom. Rev.* **1990,** *9* (1), 37-70.
22. Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R., PRINCIPLES AND PRACTICE OF ELECTROSPRAY IONIZATION - MASS-SPECTROMETRY FOR LARGE POLYPEPTIDES AND PROTEINS. *Mass Spectrom. Rev.* **1991,** *10* (5), 359-451.
23. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998,** *17* (1), 1-35.
24. Oepts, D.; Van der Meer, A. F. G.; Van Amersfoort, P. W., THE FREE-ELECTRON-LASER USER FACILITY FELIX. *Infrared Phys. Technol.* **1995,** *36* (1), 297-308.
25. Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T., Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions. *Rev. Sci. Instrum.* **2005,** *76* (2).

26. Polfer, N. C.; Oomens, J., Reaction products in mass spectrometry elucidated with infrared spectroscopy. *Phys. Chem. Chem. Phys.* **2007,** *9* (29), 3804-3817.
27. A.G. Marshall, T. C. L. W., T.L. Ricca,, Tailored excitation for Fourier transform ion cyclotron mass spectrometry. *J. Am. Chem. Soc.* **1985**, 7893–7897.
28. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., THE DEVELOPMENT AND USE OF QUANTUM-MECHANICAL MOLECULAR-MODELS .76. AM1 - A NEW GENERAL-PURPOSE QUANTUM-MECHANICAL MOLECULAR-MODEL. *J. Am. Chem. Soc.* **1985,** *107* (13), 3902-3909.
29. Grimme, S., Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006,** *27* (15), 1787-1799.
30. Schafer, A.; Huber, C.; Ahlrichs, R., FULLY OPTIMIZED CONTRACTED GAUSSIAN-BASIS SETS OF TRIPLE ZETA VALENCE QUALITY FOR ATOMS LI TO KR. *J. Chem. Phys.* **1994,** *100* (8), 5829-5835.
31. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09. Gaussian Inc.: Wallingford CT, 2009.
32. Fales, B. S.; Fujamade, N. O.; Nei, Y. W.; Oomens, J.; Rodgers, M. T., Infrared Multiple Photon Dissociation Action Spectroscopy and Theoretical Studies of Diethyl Phosphate Complexes: Effects of Protonation and Sodium Cationization on Structure. *J. Am. Soc. Mass Spectrom.* **2011,** *22* (1), 81-92.
33. Tesmer, J. J. G.; Sunahara, R. K.; Johnson, R. A.; Gosselin, G.; Gilman, A. G.; Sprang, S. R., Two-metal-ion catalysis in adenylyl cyclase. *Science* **1999,** *285* (5428), 756-760.
34. Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J., STRUCTURES OF TERNARY COMPLEXES OF RAT DNA-POLYMERASE-BETA, A DNA TEMPLATE-PRIMER, AND DDCTP. *Science* **1994,** *264* (5167), 1891-1903.
35. Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J., Crystal structures of human DNA polymerase beta complexed with DNA: Implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* **1996,** *35* (39), 12742-12761.
36. Arndt, J. W.; Gong, W. M.; Zhong, X. J.; Showalter, A. K.; Liu, J.; Dunlap, C. A.; Lin, Z.; Paxson, C.; Tsai, M. D.; Chan, M. K., Insight into the catalytic mechanism of DNA polymerase beta: Structures of intermediate complexes. *Biochemistry* **2001,** *40* (18), 5368-5375.

37. Klahn, M.; Mathias, G.; Kotting, C.; Nonella, M.; Schlitter, J.; Gerwert, K.; Tavan, P., IR spectra of phosphate ions in aqueous solution: Predictions of a DFT/MM approach compared with observations. *J. Phys. Chem. A* **2004,** *108* (29), 6186-6194.
38. Nakanishi, K.; Hashimoto, A.; Pan, T.; Kanou, M.; Kameoka, T., Mid-infrared spectroscopic measurement of ionic dissociative materials in the metabolic pathway. *Appl. Spectrosc.* **2003,** *57* (12), 1510-1516.
39. Wang, L. J.; Yang, L. G.; Keiderling, T. A., VIBRATIONAL CIRCULAR-DICHROISM OF A-FORM, B-FORM, AND Z-FORM NUCLEIC-ACIDS IN THE PO(2)(-)STRETCHING REGION. *Biophys. J.* **1994,** *67* (6), 2460-2467.

III. Molecular Structure and Electronic State Dynamics of Isolated Thymine and Uracil

Introduction

The chromaphores associated with DNA and RNA are comprised of two types; purines and pyrimidines of which guanine and adenine are of the first type and cytosine, thymine and uracil of are the second type (Figure 13). The dissipation of energy after UV absorption of the bare nucleobases is on the order of, ~ 10^{-12} s, thus experiments have a difficult time unraveling the excited state dynamics involved. To learn more about the ultrafast decay mechanisms a choice must be made to study them in the time or frequency domain.

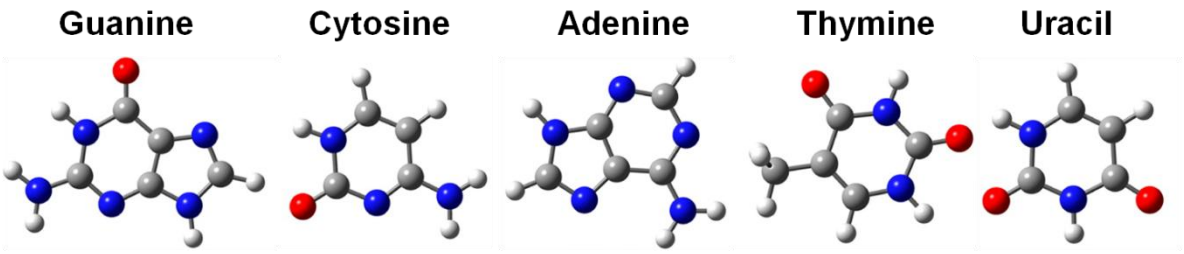


Figure 13. Ground state structures for all five nucleobases of DNA and RNA.

Thymine and uracil exhibit markedly different photo-physics from the other nucleobases. Understanding these differences is of great importance due to the dimerization of adjacent thymines in single and double stranded DNA[1] which can led cancers such as melanoma and carcinoma. Experiments measuring triplet state formation and dimer formation together have concluded that the cyclobutane pyrimidine dimer is most likely formed from a vibrationally excited ground state in thymine and a triplet excited state in the

adjacent thymine.[2,3] Excited triplet states have been identified in solution that can live for nanoseconds or longer and are populated in picoseconds after UV excitation.[4,5] They are also predicted in theoretical models to be populated by multiple pathways after UV excitation.[6,7,8] Many calculations predict these molecules exhibit excited state dynamics that have common as well as unique internal conversion (IC) and intersystem crossing (ISC) pathways.[9,10,11,12] Even with protective ultrafast IC pathways a small percentage thymine and uracil decay to a long lived dark state where they can adversely react in DNA to form covalently bonded dimers.[13]

Calculations estimating excitation energies and excited state dynamics show a large geometry change between the S_2 ($^1\pi\pi^*$) excited state and the ground state as well as multiple internal conversion and intersystem crossings from the initially excited S_2 ($^1\pi\pi^*$) state which can lead to either a hot ground state, S_1 ($^1n\pi^*$) state, T_2 ($^3n\pi^*$) or T_1 ($^3\pi\pi^*$) states.[6,7-8,11,14] Dynamics calculations estimating the decay times and state populations show efficient decay from S_2 ($^1\pi\pi^*$) through S_1 ($^1n\pi^*$) back to the ground state on the picosecond timescale. These calculations are predicting the ultrafast processes but are missing the nanosecond component. Although no dynamics calculations have estimated singlet-triplet crossing rates, two separate calculations predicting spin-orbit coupling constants for thymine are in agreement that the long lived component is a triplet state and most likely populated through the intermediate S_1 ($^1n\pi^*$) state. A different decay leading to the triplet state for uracil is purposed by Climent et al.[8] were the S_2 ($^1\pi\pi^*$) crosses directly to the T_2 ($^3n\pi^*$) then decays to T_1 ($^3\pi\pi^*$) where it becomes trapped due to small spin-orbit coupling between T_1 and the ground state. A recent calculation using TD-DFT has estimated cyclobutane formation through stacked thymines that is energetically accessible from the T_1 ($^3\pi\pi^*$) state which like

the other thymine calculations is most efficiently populated through the intermediate S_1 ($^1n\pi^*$).[15]

The photo physics of thymine and uracil have been studied extensively in solution[16] and the gas phase. The first high resolution gas phase UV spectrum for both molecules was published by Brady et al. and was the first UV spectrum of nucleobases in a molecular beam.[17] The spectrum for both molecules shows a broad absorbance over the 1000 cm^{-1} range. This was attributed to mixing of electronic states and a large geometry change upon excitation. As mentioned above, calculations of the ground and excited state structures for these molecules show that this is a likely the reason for the broad spectrum. More recently 1-methylthymine, 1-methyluracil and their water clusters have been studied in molecular beams using resonant two photon ionization (R2PI), IR-UV double resonance and UV-UV pump-probe spectroscopy.[18] The R2PI spectra for these molecules and clusters are broad, similarly to thymine and uracil, as well they all have long lived states of 100 ns or greater. The lifetime measurements reveal the long lived state survives hydration and has a longer lifetime with hydration. 1-methylthymine has since been revisited by IR-UV double resonance and UV-UV pump-probe spectroscopy. This experiment revealed that the >300 ns long lived dark state for 1-methylthymine is triplet in character. This assignment was made by comparison of the excited state IR spectrum with calculated frequencies.[19] The UV spectrum of thymine and uracil has also been measured by electron energy loss spectroscopy[20,21] and thymine by REMPI spectroscopy.[22] Neither of these experiments provide spectra which resolve the vibronic contributions thus the details of the UV spectrum and how it relates to their dynamics was not revealed but valuable excitation energies relating to the dark states were obtained from the electron energy loss spectra.

Ultrafast pump-probe experiments have identified multiple decay time constants that range from femtoseconds to nanoseconds.[23,19,22] Conclusions about the number of decays and which decay times correspond with which pathways give an inconsistent picture of the dynamics. Ultrafast experiments typically excite at fixed energies well above the excitation threshold and the broadband nature may excite many electronic states and possible tautomer contributions, thus unraveling the multi-exponential decay is not straight forward. In cases such as thymine and uracil with many decay pathways over a long time range conflicting results ensue.

Here we present data for thymine and uracil that includes high resolution UV spectra, IR-UV spectra of the ground and excited states and wavelength dependent lifetime measurements in the nanosecond to microsecond time range. Comparison of the data to past experiments and calculations suggest a revision on the current view of thymine and uracil's UV absorption and dynamics involving the hot ground states and triplet states.

Methods

A combination of molecular beam laser spectroscopy and mass spectrometry provides a wealth of detailed information that is obtainable by few other techniques. When combined with laser desorption a large range of non-thermally volatile molecules can be studied and a better understanding of their electronic state properties, molecular dynamics and structure is revealed. A detailed description of both the instrument required for such experiments as well as various spectroscopic techniques used will be given here. The various parts of the instrument, shown in Figure 14, will be described first followed by multiphoton laser spectroscopy techniques.

1. Instrumentation

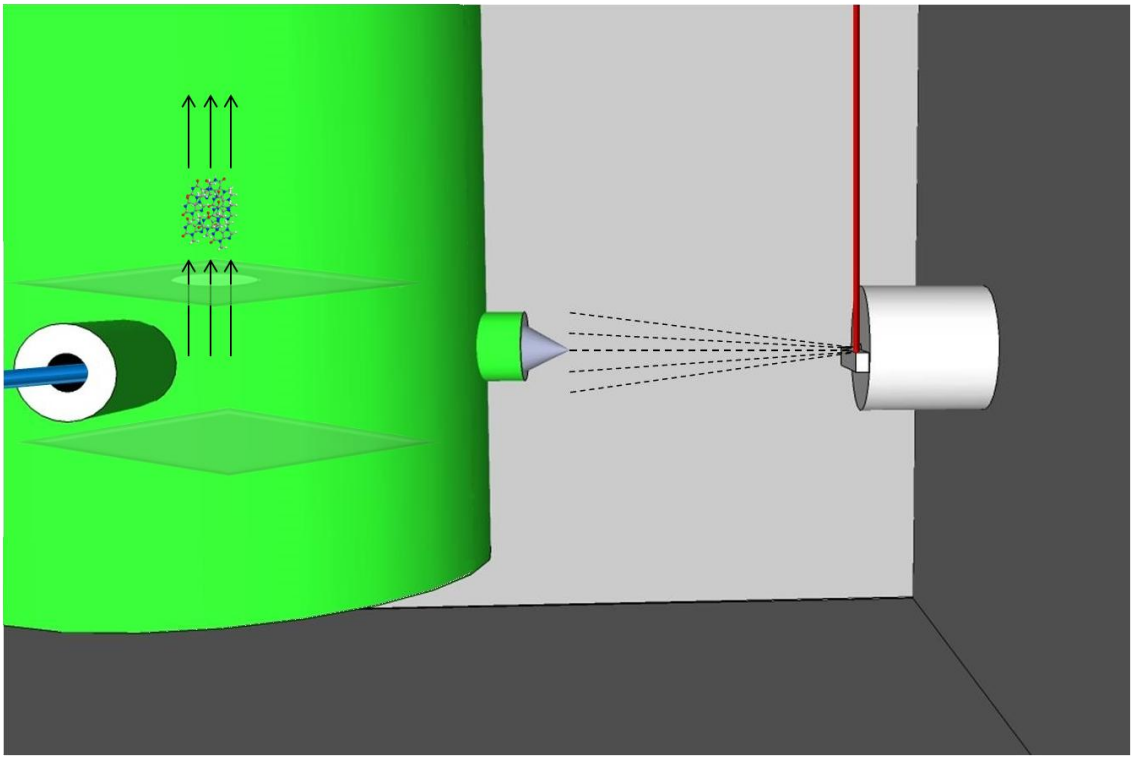


Figure 14. Inside of instrument showing laser desorption (red line), followed by molecular beam entrainment (dotted line) and photo-ionization (blue line) leading to mass analysis (triple black arrows)

1.1 Laser Desorption

Prior to the advent of laser desorption (LD) and electrospray ionization (ESI), gas phase experiments were limited to thermally labile molecules. Thus the study of biologically relevant molecules was almost completely limited to solution and solid phase. In the early 1960's the first experiments emerged that studied the formation of ions from surfaces by laser irradiation. By the mid 1980's Large molecular ions were being studied in the gas phase by a handful of researchers. These studies combined laser desorption with mass spectrometry to study wavelength and desorption substrate dependence on molecular ion

production.[24,25] Fundamental experiments such as these have lead to our current understanding of the laser desorption mechanism. A range of substrates and laser wavelengths have been identified which optimize the vaporization of target analytes while at the same time minimizing fragmentation. The complete mechanism for laser desorption has not been fully characterized but certain aspects of the process have emerged from experiments.[26] The substrate or matrix used for desorption of the analyte serves as a mediator for energy absorption. Heating the substrate at sufficiently high rates, $> 10^{11}$ K/s, favors vaporization of analytes over chemical degradation.

The desorption laser used for the following experiments is a Nd:YAG laser, 10 ns/pulse, with both second and fourth harmonic options. This allows for the selection of wavelengths; 1064 nm, 532 nm and 266 nm. For typical non-volatile organic molecules a graphite substrate is used with 1064 nm desorption (~1 mJ/cm^2). This combination of substrate, wavelength and energy was found to give the best shot to shot signal average while at the same time maintaining high ion signal. Other substrates and wavelengths will be discussed in the experimental section of chapter 4.

1.2 Molecular beam cooling

Gas phase molecules of low internal energy can be prepared by a process known as jet cooling. This occurs through collisions between laser desorbed analytes and a molecular beam drive gas. Noble gases are the best choice of drive gas for cooling molecules in a supersonic jet due to only having translational degrees of freedom in which to store energy. Argon is the gas used in these experiments (4500 torr backing pressure) as it previously has been shown cool mid to large polyatomic molecules to low vibrational and rotational

energy.[27] The molecular beam of Argon is timed to pulse promptly after laser desorption. As the argon expands from the high pressure region, ~ 4500 torr, to the low pressure region, 2×10^{-5} torr, through the molecular beam nozzle the mean velocity increases while the velocity distribution narrows as shown in Figure 15.[28] The narrowing of the velocity distribution

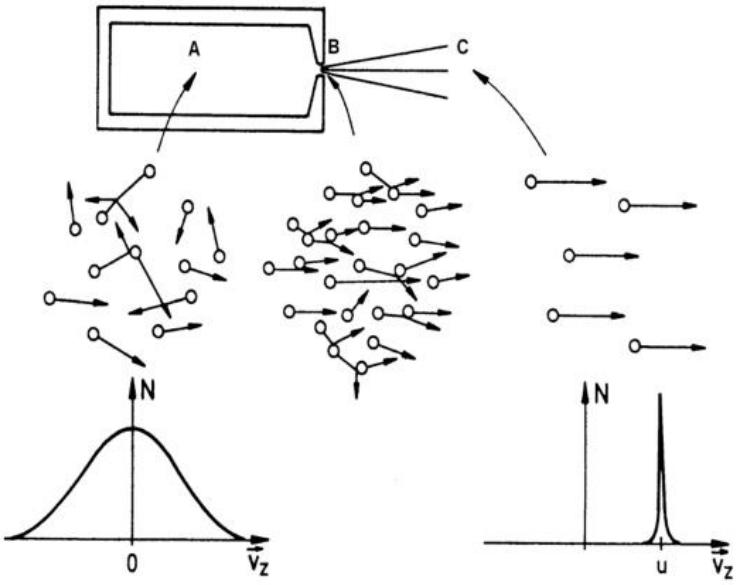


Figure 15. (A) Velocity distribution and mean speed inside molecular beam valve. (B) Distribution of velocity vectors at the molecular beam exit. (C) Velocity distribution and mean speed in expanded molecular beam

occurs through removal of translational degrees of freedom which corresponds directly to translational temperature. As the molecular drive gas collides with the laser desorbed molecules, translational and thermal energy equilibration takes place removing rotational and vibrational energy from the molecules.[29] This process leads to key features of molecular beam spectroscopy. The molecules are cooled to their ground vibrational state leading to

vibronically resolved UV spectra with rotational envelopes below 20 K.[27] During perturbation the molecules are isolated, free of collisions and solvent interactions so the spectra can be directly compared to ab initio calculations. Lastly conditions such as backing pressure and choice of drive gas can create clusters of the analyte molecules and or drive gas so that intermolecular interactions can be studied.

1.3 Time of Flight Mass Spectrometry

Mass spectrometry is an attractive method for identification of ions produced in molecular beams due to its inherent ability to identify different molecular species by mass and the short time frame for detection. Time of flight mass spectrometry (TOF-MS) is well suited for coupling with laser spectroscopy. The method is based on identifying ions of a particular mass to charge ratio by measuring the time it takes the ions to travel through a field free region after being accelerated through a region of electrical potential.

$$t = \frac{L}{v} = L\sqrt{\frac{m}{2qV}} \qquad 1$$

In equation 1, t is the time it takes ions to travel from the ionization source to the detector, L is the length of the flight path, m/q is the mass to charge ratio and V is the potential difference between the repeller and extraction electrodes (See Figure 11). There are two advantages of TOF-MS when coupled with laser spectroscopy. The detection of ions over a large mass range occurs within 50-100 μs thus ions produced with excess internal energy have little time to fragment before they are detected. As well the ionization event occurs in only a few nanoseconds thus defining time zero of the TOF quite well. Although one disadvantage to using lasers for ionization is that the ions created have a spatial distribution mirroring the laser spot size. A spread of initial locations in the potential region leads to a

spread in time for any m/z ion packet arriving at the detector. This in turn limits resolution, see eqn. 2. To alleviate this a reflectron, or electrostatic mirror, can be used to compensate for the spread in energy of the ion packets. It consists of a series of ring electrodes with increasing potential that are placed in the drift region (Figure 16). Ions with greater kinetic energy travel farther into the reflectron before they are sent back out to the detector. So by tuning the electrical potential of the reflectron, ions of a given m/z can be focused in time at the detector thus correcting for the distribution in kinetic energy and alleviating the mass resolution problem. Essentially the reflectron focuses ion packets in time on the detector, as shown in equation 2, creating a shorter arrival time distribution and thus better mass resolution.

$$Mass\ Resolution = \frac{t}{\Delta t} = \frac{m}{\Delta m} \qquad 2$$

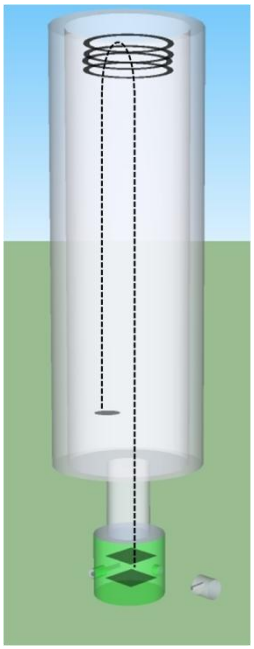


Figure 16. Ion trajectories in a reflectron time of flight mass spectrometer

2. Laser Spectroscopy

Spectroscopy is the study of the interaction of radiation with matter. When measuring absorbance in the gas phase a traditional spectrometer lacks the molecular density required to measure a difference in radiation before and after molecular interaction.

$$A = \sigma \times c \times L \qquad 3$$

The Beer-Lambert law shown in eqn. 3 defines spectroscopic measurements by the absorbance A, the absorbance cross-section σ, the molecular concentration c, and L the path length for which the radiation travels through the sample. For studying molecules in a molecular beam two important changes to traditional spectroscopy must be made. The use of a high intensity light source, such as lasers, ensures multiple photon molecule interactions despite low molecular densities. As well the use of a detection method other than direct absorbance such as fluorescence emission, fragmentation or ionization greatly enhances

one's ability to measure the interaction. These alternate forms of detection are termed "action" spectroscopy because they are not measuring the true absorbance of a molecule but rather some effect due to absorbance. When coupling laser desorption, molecular beam jet cooling and resonance enhanced multiphoton ionization laser spectroscopy the problems of low signal to noise in traditional gas phase spectroscopy can be overcome.

2.1 Resonance Enhanced Multi-photon Ionization Spectroscopy

Resonance enhanced multi-photon ionization spectroscopy (REMPI) is a technique which enhances traditional laser spectroscopy by interacting with multiple states of a molecule in which the last state created from the overall interaction is the parent ion of the molecule under study. An example of an intermediate electronic state interaction is given in Figure 14. This technique is favorable for gas phase spectroscopy because mass spectrometers can detect the ionized molecule quite sensitively so the interaction of photons with just a few molecules can be measured. As well this gives us the ability to study the interaction of UV and IR radiation with molecules in a mass selective manner. Many organic and biological molecules have ionization potentials (IP) between 7.5 and 10.5 eV thus two UV photons are required for ionization. One photon excites an electronic state and the second ionizes the molecule. This is termed one color two photon ionization. If the excited electronic state is energetically less than half of the IP than a second laser with higher photon energy is required and is termed two color two photon ionization. In either case the first laser is scanned in the frequency domain while ionization is monitored in a mass spectrometer and a mass specific UV action spectrum is collected. A Jablonsky diagram depicting vibronic excitation is shown in Figure 17. In the case of jet cooled spectroscopy the resolved UV spectrum consists of multiple isomers, confirmations and vibronic

contributions from each mass being monitored. As will be discussed below this turns out to be the key to studying how different isomers, tautomers and confirmers interact with UV radiation.

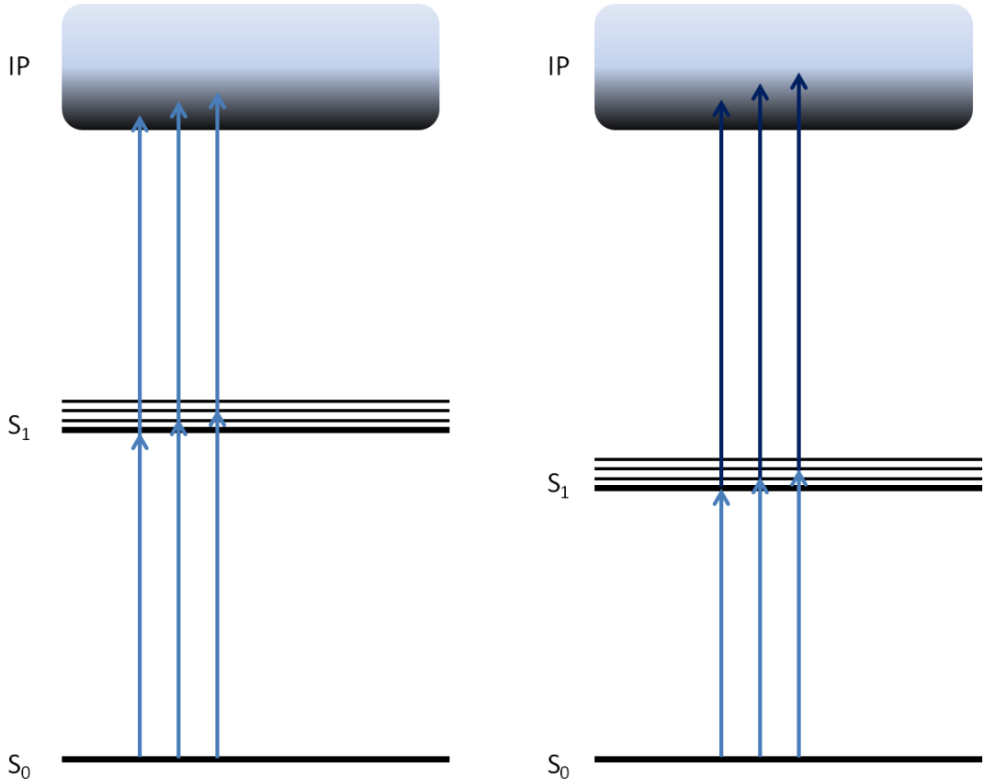


Figure 17. Jablonsky diagram for one color and two color REMPI

2.2 IR-UV Double Resonance Spectroscopy

Due to the vibrational temperature of molecules seeded in a molecular beam the UV spectrum is highly resolved. The vibronic transitions originate from the ground electronic and ground vibrational state. Absorbance of an IR photon prior to UV excitation removes molecules from the ground vibrational state and thus depletes the corresponding resonant UV transition as depicted in Figure 18.

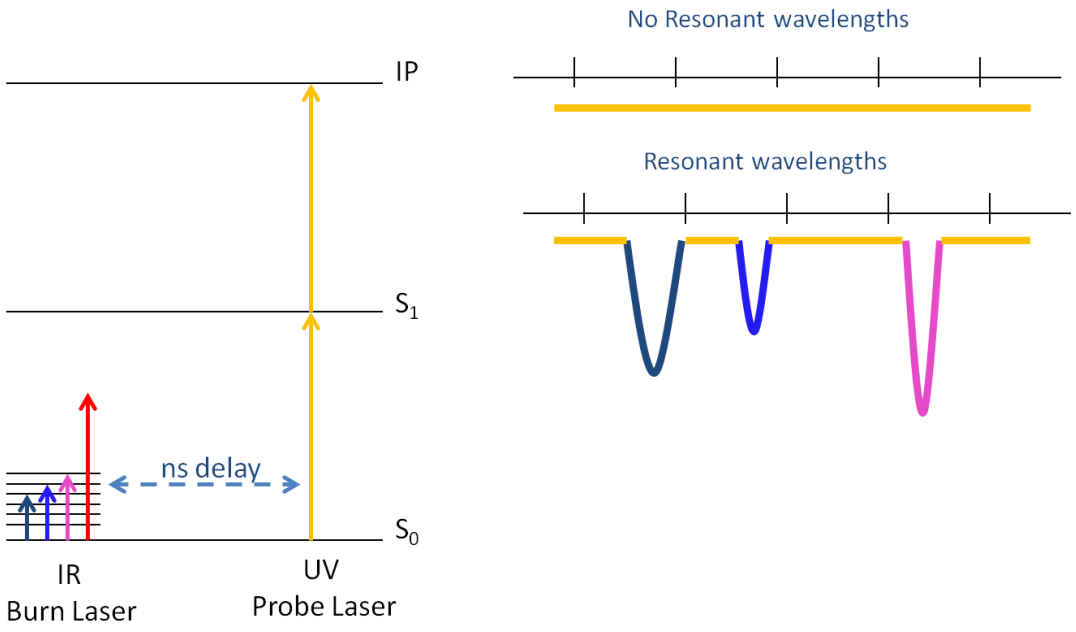


Figure 18. (Left) Jablonsky diagram of IR-UV double resonance spectroscopy. (Right) simulated dip spectrum.

In this manor an IR spectrum can be collected for each UV transition in the R2PI spectrum. A somewhat special case of IR-UV spectroscopy was used to collect the IR spectrum for thymine and uracil. These molecules have unique UV absorption in the sense that they absorb UV photons nearly continuously over a broad range and depletion of the UV absorbance does not occur from IR excitation of the ground state. Another approach was used to collect the ground state IR spectrum. The UV laser was set to a frequency lower in energy then the onset of the UV spectrum, where no ionization could be seen by the REMPI laser. 30 ns prior to firing the UV the IR was scanned over the region of 3200 - 3700 cm^{-1} and a gain in ionization was measured as a function of the IR frequency. The excited state IR spectrum can also collected in some cases by delaying the IR laser 10 ns after the UV

excitation and then delaying the probe UV laser by another 30 ns after the IR pulse. Here a depletion in the ion signal was measured.

2.3 UV-UV Pump-Probe Spectroscopy

Resonant electronic states measured in two color REMPI experiments can be studied in more detail by delaying the ionization laser in time after UV excitation. When using nanosecond lasers the experiment is limited to studying dynamics on this time scale but the pyrimidine nucleobases exhibit nanosecond or longer excited state lifetimes from a small population of the UV excited molecules and these longer lived electronic states can be studied in detail.

3. Experimental

Using the instrument described in section one the two color R2PI spectrum for thymine and uracil was collected. The standards were purchased from Sigma Aldrich and analyzed without further purification. The samples are prepared by mixing the standard with ground graphite, this is freshly prepared for each sample by roughing the surface of the graphite sample bar with fine grit sand paper, then rubbing the sample onto the roughened surface. The bar is then placed on a translating stage inside the high vacuum chamber directly in front of a pulsed molecular beam valve. Laser desorption is accomplished by 1064 nm light from a Continuum minilight Nd:YAG laser (15 mJ/pulse). The laser is attenuated to 1 mJ/pulse or lower and focused on the sample at the exit of the molecular beam nozzle. After desorption the sample molecules are entrained in a molecular beam of Argon, 6 atm backing pressure, where they travel to the ionization region of the instrument. The neutral molecules are ionized by resonance enhanced multi-photon ionization (REMPI) and the subsequent

ions are detected by a reflectron time of flight mass spectrometer. The first and resonant photon comes from the frequency doubled output of a lumonics HD-300 tunable dye laser pumped by either the second or third harmonic of a quanta ray DCR 2A. The maximum spectral line width and pulse energy are 0.4 cm^{-1}, 0.3-0.7 mJ respectively. The second photon comes from a GAM EX-5 excimer laser firing at 193 nm. (1.5-2 mJ/pulse). The excimer laser spot size is cut down with an iris to the same size as the dye laser (0.5 cm dia) and the co-propagating beams are checked for spatial and temporal overlap prior to each wavelength scan. For IR-UV double resonance spectroscopy an OPO/OPA (Laservision) is pumped by the fundamental of a quanta ray DCR 2A (< 550 mJ/pulse) giving a tunable output of 2800 cm^{-1} to 3800 cm^{-1} with 3-5 mJ/pulse over the range and has a maximum spectral line width of 3 cm^{-1}. The IR laser is focused in the ionization region and counter-propagates the two UV lasers. Laser overlap and frequency calibration are preformed on previously studied molecules.

4. Computational

The minimal structure, harmonic and anharmonic frequencies of the ground and excited states were calculated at the Moller-Plesset perturation theory[30] (MP2) level and with the second-order algebraic diagrammatic construction (ADC(2)) method[31,32] with the resolution of identity.[33] The cc-pVTZ basis set[34] was used throughout the calculations. Anharmonic frequency calculations at the second-order perturbation theory (PT2) level were carried out for the ground electronic state using Gaussian09.[35] One-dimensional NH stretching frequencies (1D-scan) were obtained by solving (1) using the Numerov-Cooley integration technique[36]

$$\left\{-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r_{NH}^2}+V(r_{NH})\right\}\chi_n(r_{NH})=E_n\chi_n(r_{NH}) \quad 4$$

where $\frac{1}{\mu}=\frac{1}{m_N}+\frac{1}{m_H}$ and $V(r_{NH})$ is a potential energy surface scan over the N-H bond coordinate. Assuming small coupling of the NH stretching motion with other vibrational degrees of freedom, the 1D-scan results should be comparable to one-dimensional anharmonic frequencies evaluated from diagonal anharmonic constants (denoted as 1D-PT2).

Results and Discussion

1. Resonance Two-Photon Ionization Spectroscopy (R2PI)

The UV spectrum for both Thymine and Uracil have been revisited after 15 years. As mentioned in the introduction the first UV spectrum for bare nucleobases, thymine and uracil in a molecular beam, was collected by Brady et al in 1988.[17] Here the same UV region is scanned at high resolution to further analyze the vibronic contributions, as well as identify possible tautomeric contributions to the spectrum, as this has not previously been identified in a laser desorption seeded molecular beam. As published previously thymine and uracil exhibit broad UV spectra (Figures 19 and 20). At 0.4 cm^{-1} resolution the averaged spectra shows narrower resonances within the broad background. The spectrum of thymine clearly shows peaks well above the signal to noise level where as for uracil the peaks are less pronounced due to lower ion signal over the range. These contributions to the spectrum appear to be vibronic transitions to the lowest lying vibrational modes but a

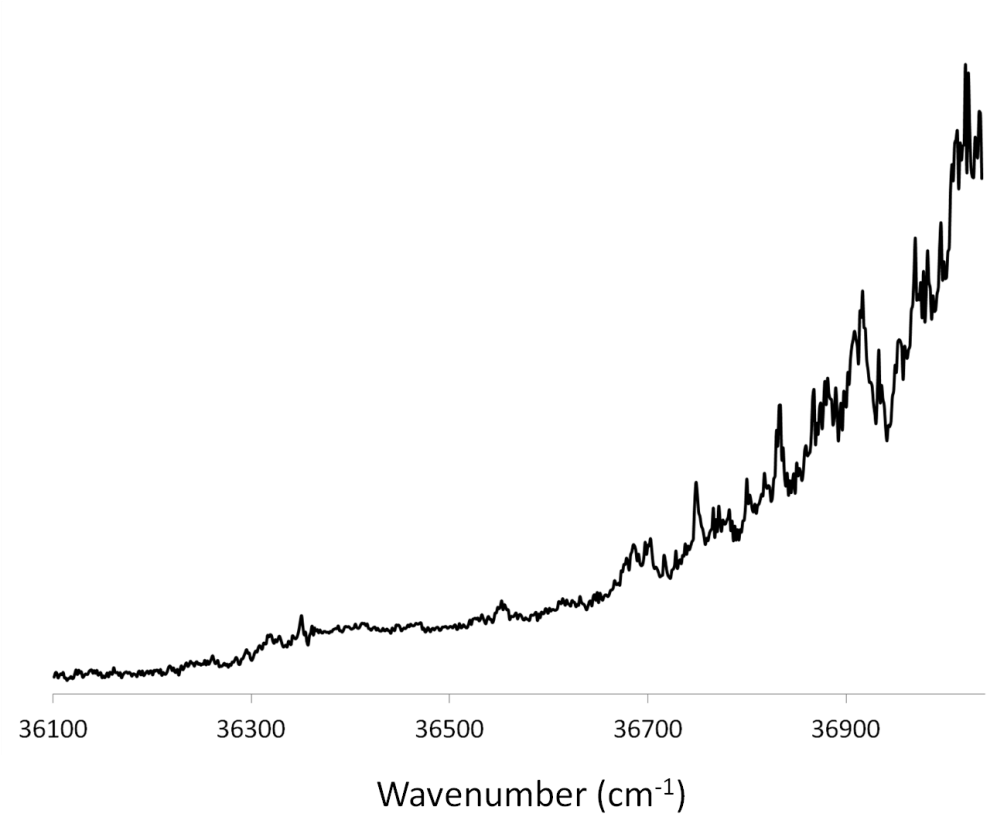


Figure 19. Onset of the UV spectrum for thymine recorded with second color at 193nm and 10 ns delay.

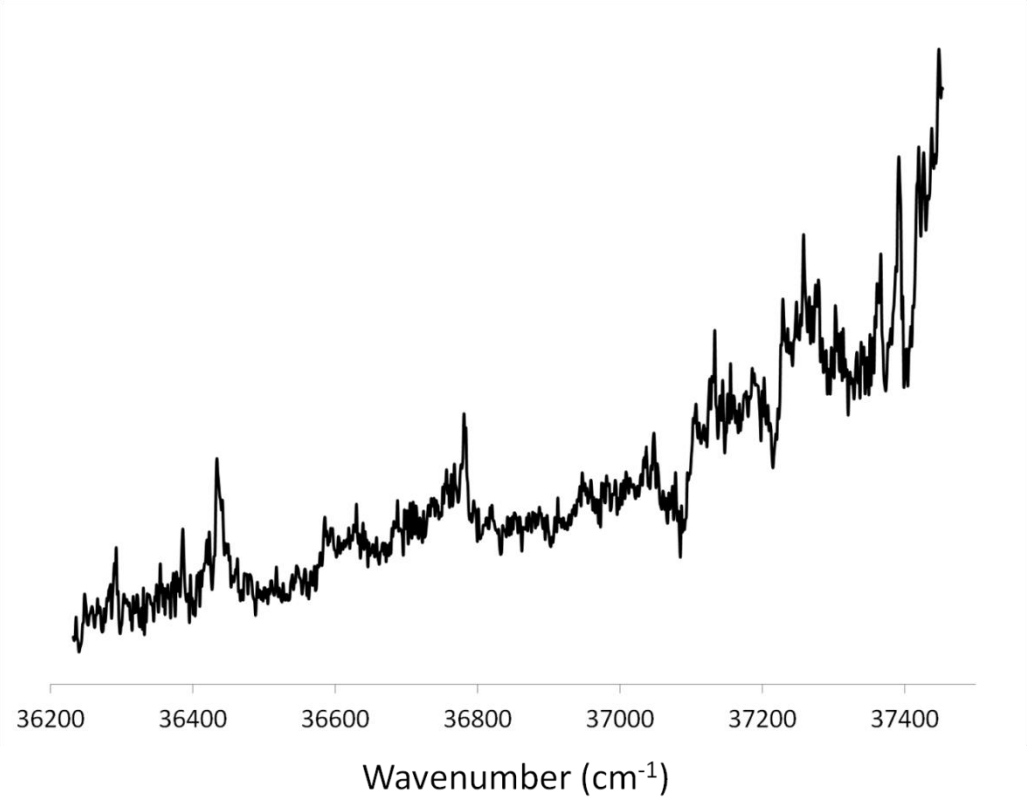


Figure 20. Onset of the UV spectrum for uracil recorded with 193nm second color and 10 ns delay.

Franck-Condon analysis using TD-DFT failed to predict any vibronic contributions between S_0 and S_2 ($^1\pi$-π^*) due to a large geometry change between the minimum energy structures. A frequency analysis using ground state harmonic frequencies was able to match many of the features seen in the UV spectrum which is shown in section 6. The match of the ground state frequencies to the vibronic spectrum is a bit unexpected. This would suggest that a hot ground state is absorbing photons with high efficiency to the ($^1\pi$-π^*) singlet state for both molecules. Hot ground states are known to be produced in both thymine and uracil through internal conversion in less than 1 picosecond after UV absorption. This process of UV absorption, IC and subsequent UV absorption would explain the broad spectrum and

assignments of vibronic structure in the UV spectrum but one would expect that molecules with ~ 4 eV of internal energy would not have narrow features as seen in these spectra. Although if the energy transfer from vibrations to rotations in the hot ground state is small the vibronic transitions from the hot ground state to the $^1(\pi\pi^*)$ could be narrow if the excitation was to a region of low vibrational state density. This would only occur if there were an unusually large change in geometry between the states or if the excitation was to a different state say for instance the $^1(n\pi^*)$. If this were the case then the REMPI spectrum on the nanosecond time scale could be monitoring vibronic contributions in the broad UV spectrum that lead preferentially to the long lived state where as vibronic contributions that are not seen could be due to unfavorable Franck-Condon overlap from the hot ground state or after excitation these geometries or starting locations on the excited potential surface could preferentially lead back to the ground state through IC. The latter has been suggested by Lobsiger et al.[37] for cytosine and is discussed in section 6. Comparison of the calculated ground/excited state structures and IC geometries with ground state vibrational frequencies hint at this possibility for thymine and uracil.

2. Ground State and Excited State IR spectra

The IR spectrum for the ground state is collected by IR-UV double resonance spectroscopy. Due to the lack of an IR-UV hole burning spectrum at multiple UV wavelengths tested, a gain spectrum was attempted where the UV laser was tuned to 35714 cm^{-1} which is just prior to the onset of UV absorption. The resulting spectrum is shown in Figures 21 and 22 where the lower trace is the ground state spectrum and the upper trace is the dark state spectrum. Two frequencies in the N-H stretch region are seen for the ground state corresponding to the diketo forms for both thymine and uracil.

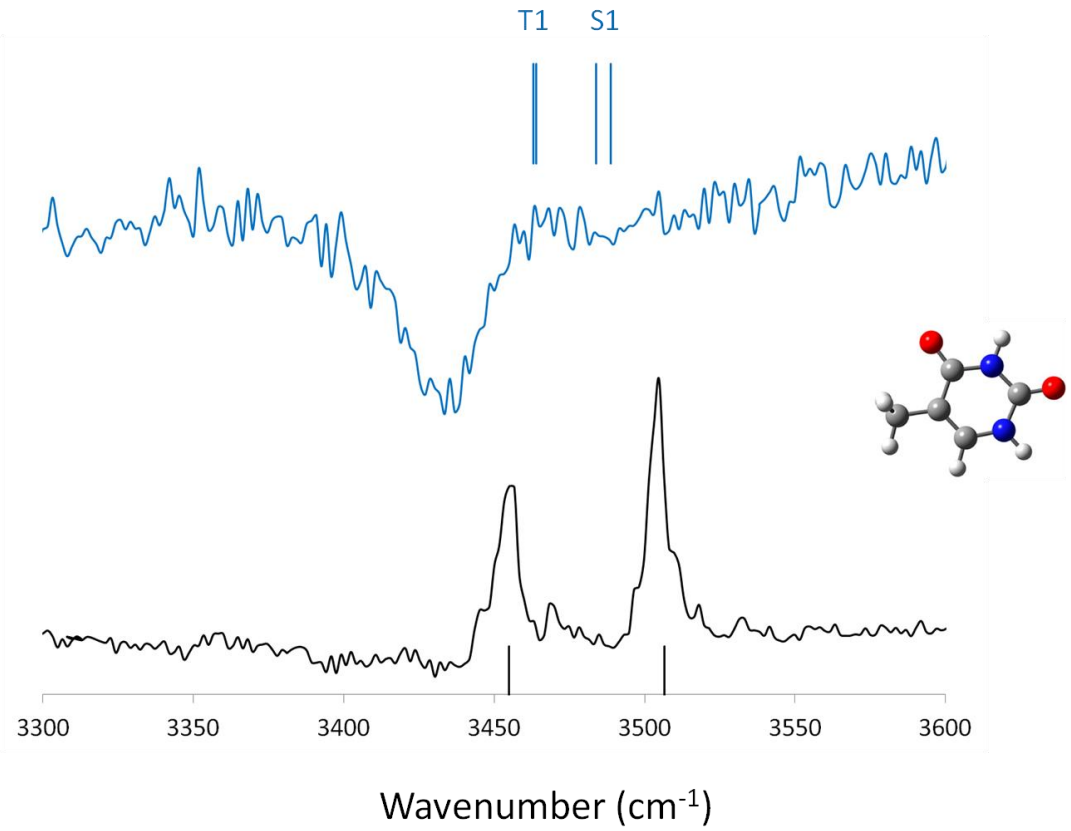


Figure 21. (Lower) ground state IR spectrum for thymine with scaled anharmonic frequency calculations for the N1-H (3455 cm^{-1}) and N3-H (3505 cm^{-1}) vibrational modes. (Upper) excited state IR spectrum with the scaled anharmonic frequencies of both the triplet and singlet excited states. Scaling factor for all frequencies is 0.991676.

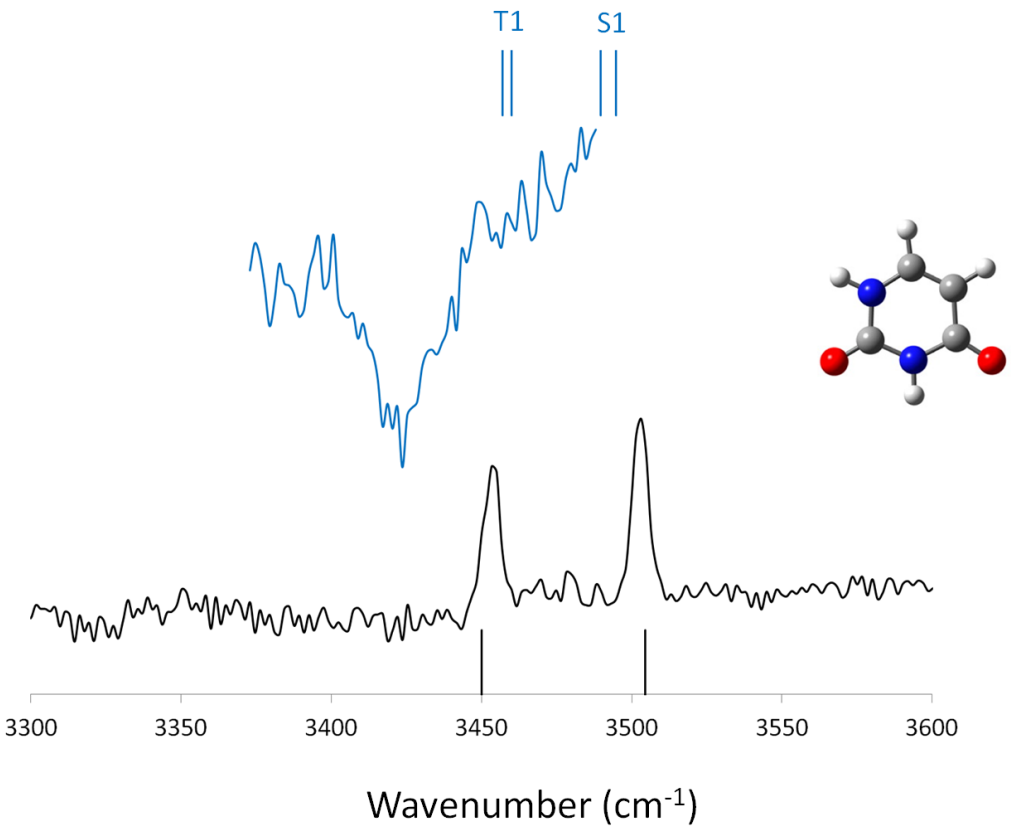


Figure 22. (Lower) ground state IR spectrum for thymine with scaled anharmonic frequency calculations for the N1-H (3452 cm^{-1}) and N3-H (3501 cm^{-1}) vibrational modes. (Upper) excited state IR spectrum with the scaled anharmonic frequencies of both the triplet and singlet excited states. Scaling factor for all frequencies is 0.991676.

The excited state spectrum shows only one peak which is red shifted and broadened in comparison to the ground state frequencies. Harmonic and anharmonic analysis of this system suggest that the two N-H stretches are coupled in both the S_1 ($^1n\pi^*$) and T_1 ($^3\pi$-π^*) excited states. In the excited states the two modes have frequencies predicted between the ground state modes making a direct assignment unclear. In the singlet state the lower energy mode, N1-H, has a large blue shift while the N3-H has a small red shift resulting in a

spectrum with a broad absorbance near the N3-H frequency of the ground state. In the triplet state the coalescence and shift is a small blue shift of the N1-H and a large red shift of the N3-H thus leaving a broad absorbance near the N1-H ground state spectrum. The experimental spectrum for both molecules is a single broad peak just to the red of both amine stretches. After ISC or IC the gain in internal energy will cause a red shift of all the vibrational frequencies. The overall shift is dependent on the difference between the excited state and ground state structures, the amount of internal energy, the anharmonicity constants and the number and type of modes in the molecule.[38] If we are to assume a similar anharmonicity for both states and we use calculated adiabatic energies for the two states then the red shift associated with the triplet state would be three times that of the singlet based on the amount of internal energy. With the ($^3\pi$-π^*) triplet state being lower in energy than the vertical excitation by ~1.5 eV while the (^{1}n-π^*) singlet is only 0.5 eV and the N-H frequencies are predicted to be 30 cm^{-1} *vs* 50 cm^{-1} to the blue of the experimental band for the triplet and singlet respectively the logical assignment of the dark state would then be the ($^3\pi$-π^*) triplet state because if the singlet N-Hs were to red shift 50 cm^{-1} with 0.5 eV internal energy then the triplet would be shifted 150 cm^{-1} with its 1.5 eV internal energy which is much too large for a molecule of this size and internal energy. A much more reasonable red shift is 30 cm^{-1} with 1.5 eV.

3. Wavelength Dependent Lifetimes of the Long Lived Dark States

Lifetime measurements with nanosecond time delays have been collected at many UV excitation wavelengths and are shown in Figures 23 and 24, plots of pump-probe decays are given in Appendix II section 1.

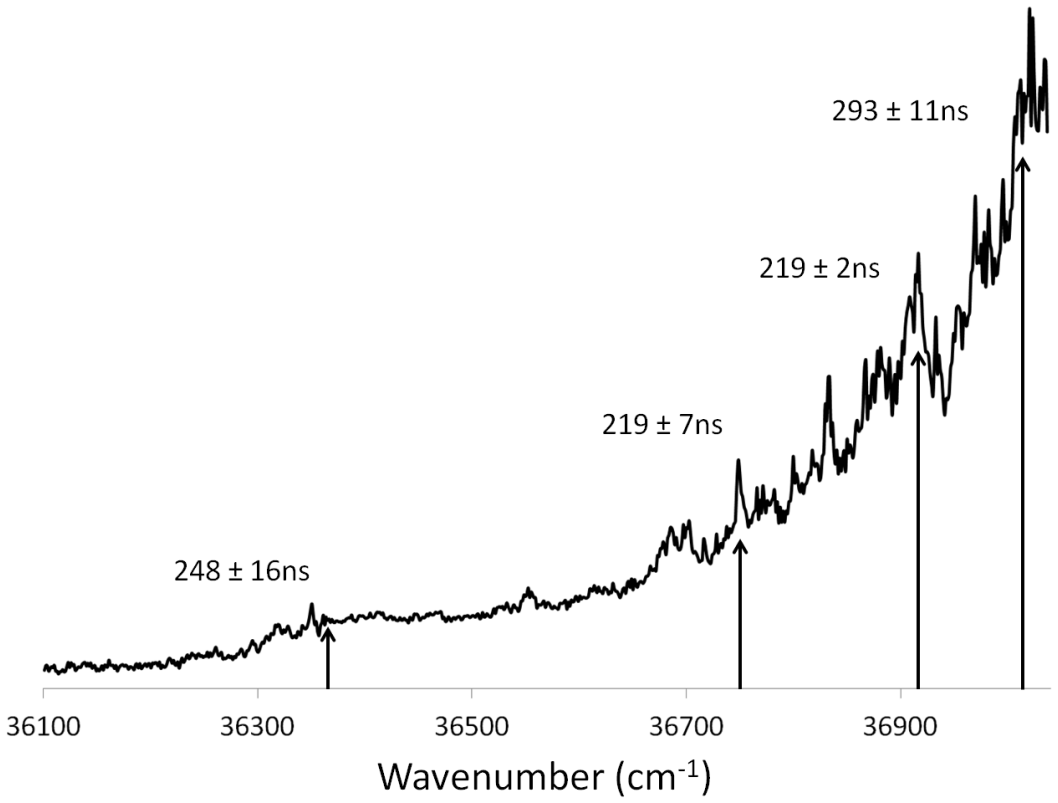


Figure 23. UV spectrum of thymine, vertical arrows are showing the locations on the spectrum for the excited state lifetime measurements. Lifetime values are given above their probe locations.

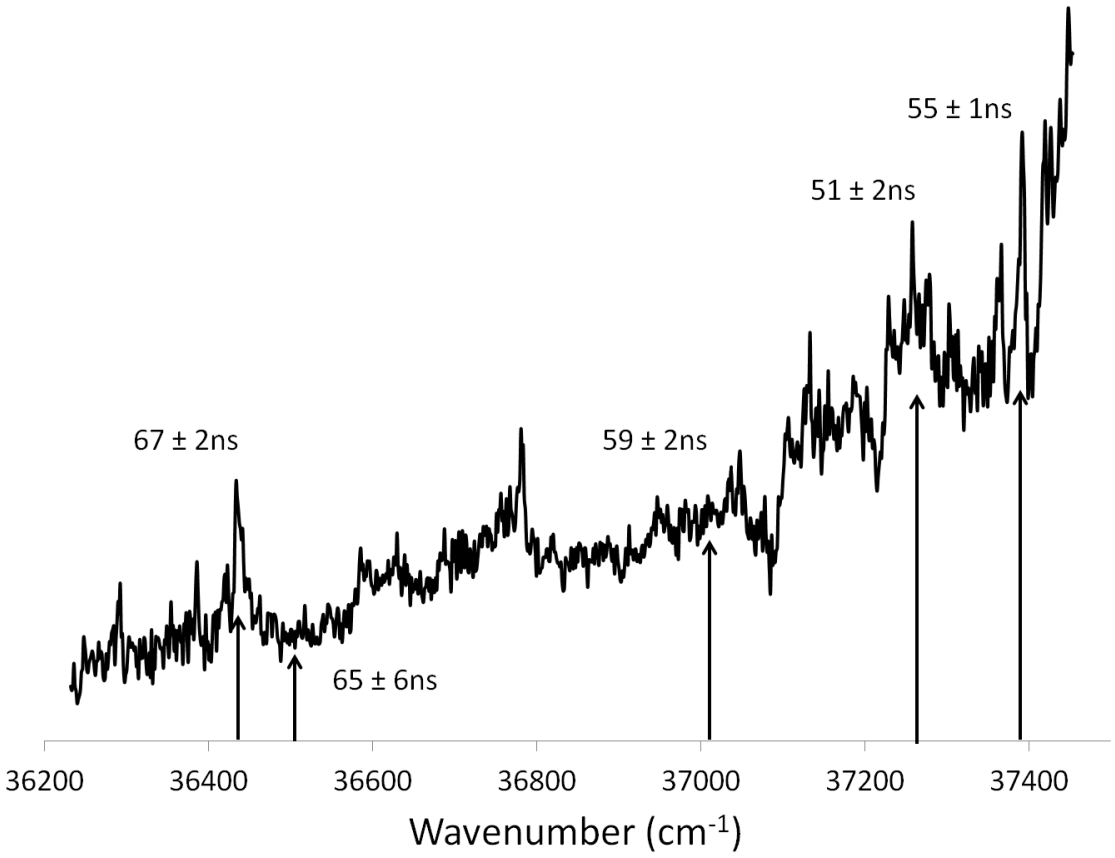


Figure 24. UV spectrum of uracil, vertical arrows are showing the locations on the spectrum for the excited state lifetime measurements. Lifetimes values are given above their probe locations.

A difference is seen between thymine and uracil both in the number of decays and in the wavelength dependence of the lifetimes. For thymine two excited state decays are measured at all pump energies except the lowest (36101 cm^{-1} shown in Appendix II) where only the nanosecond component is seen. The longer lived, > 8 μs, component has not been previously reported in other gas phase lifetimes measurements for thymine or thymine derivatives. The longest lifetime for thymine (293 ns) is measured at highest pump energy and the shortest lifetimes (219 ns) are measured at two vibronic features in the middle of the spectrum.

Uracil exhibits more common dynamics where the lifetime increases with excitation energy ranging from 50 to 70 ns. This difference in thymine and uracil is intriguing simply because calculations suggest that their dynamics should be very similar. Excitation of vibronic transitions in thymine seem to be leading to shorter decays in the long lived state but without measuring lifetimes near these peaks it is tough to tell for sure what is being measured. The presence of two decay times in thymine shows that at least two relaxation pathways are present. We propose that for thymine in the nanosecond time regime we are measuring the dark state, > 200 ns, and the hot ground state, > 8 μs, which is the elevated background at all times after 0. This state becomes populated from the ultrafast IC and two photons at 193 nm ionize the molecule through some unknown excited state. It should be mentioned that the excimer laser (193 nm) did not ionize thymine when hot ground states were prepared by the IR laser alone but only with hot ground states prepared through IC from UV. This leaves the small possibility that the > 8 μs lifetime belongs to another triplet state with very low population.

4. Onset of Ionization at 1.5 μs delay

To further interpret the time dependence of the UV spectrum for thymine a second UV spectrum was obtained at a 1.5 μs delay between UV pump and probe lasers. The onset of ionization from the microsecond component was identified and is shown in Figure 25.

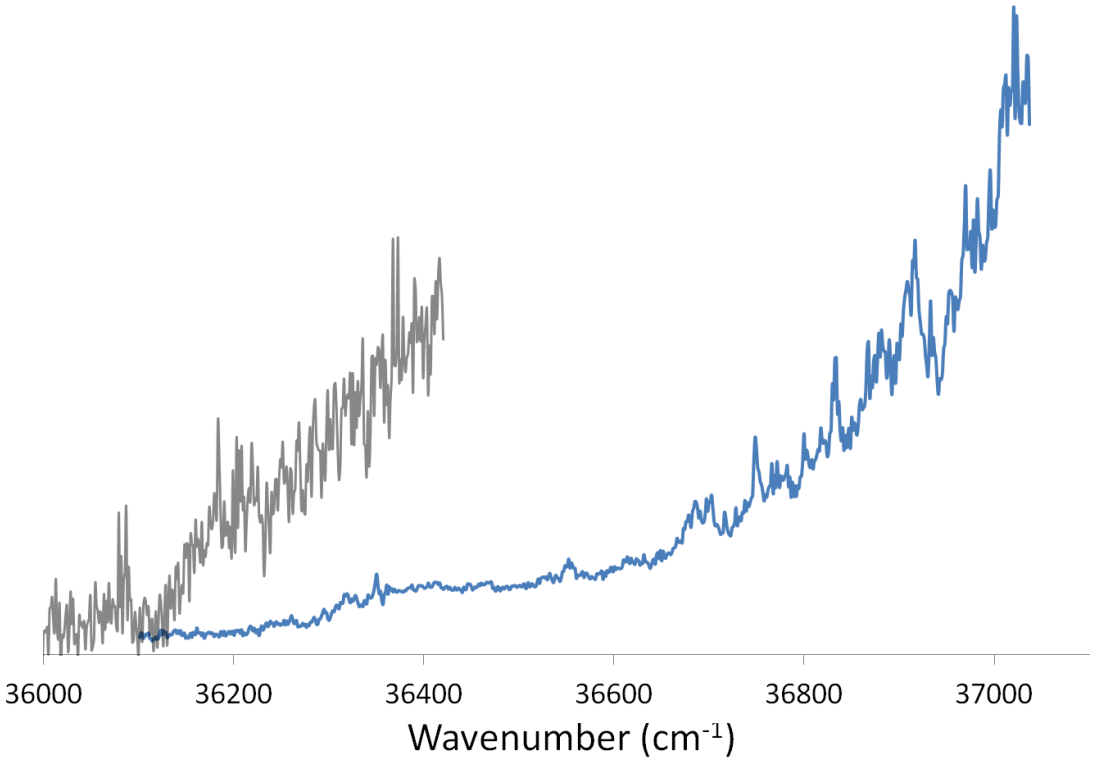


Figure 25. (Blue) UV spectrum at 10 ns probe delay. (Grey) UV spectrum, scaled 10 times in comparison to blue spectrum, collected at 1.5 µs probe delay showing the onset of ionization for the longest lived state.

The state that lives for > 8 µs, measurement limited by the instrument, has an onset energy for population at 36100 cm^{-1}. Although it is not shown in the 10 ns time delay UV spectrum the onset of ionization is ~36000 cm^{-1}. The onset of ionization at 1.5 µs delay is ~100 cm^{-1} to the blue. If this state is the hot ground state populated through IC the 100 cm^{-1} energy difference would correspond with a small barrier from the lowest energy vertical FC point to the S_2/S_0 CI. Many calculations suggest that the IC is barrier less but at this low

excitation energy we could be seeing the small barrier from the S_2 minimum to the S_2/S_0 CI predicted by linearly interpolated coordinate calculations.[9]

5. Ground State IR Pumping Followed by Excited State Lifetime Measurements

Excitation of the ground state N1-H vibrational mode (3504.7 cm^{-1}) followed by UV excitation and variable delay ionization measures excited state lifetimes while sampling a different location/geometry of the excited potential surface. The highest and lowest UV energies from the previous lifetime measurements were chosen to test the vibrational energy dependence on the lifetime of the dark state of thymine. Figures 26 and 27 show the lifetimes with (red) and without (blue) ground state IR excitation. At both wavelengths there is an increase in ion signal with UV excitation from the hot ground state and significantly shorter lifetimes of the dark state, 164 ± 5 ns at 37021 cm^{-1} and 167 ± 3 ns at 36364 cm^{-1} compared to 293 ± 11 ns and 248 ± 8 ns respectively. The change in lifetime is significant and can be explained by the following argument. After IR excitation of the ground state enough time elapses before the UV pump for complete intra molecular vibrational relaxation (IVR) to occur. When the UV pump excites this statistical ensemble of states with 3500 cm^{-1} of internal energy a vibronic transition to an excited state, presumably S_2 ($^1\pi$-π^*), but possibly S_1 (^{1}n-π^*) or higher energy excited states, occurs and subsequent decay through the dark state is faster for two possible reasons. The excitation is simply higher in energy on the potential surface and more internal energy is available to find internal conversion and intersystem crossing geometries faster or the excitation is to an unknown state from which decay through the dark state happens faster. It does not however compare to the previous lifetime reported by Kunitski et. al.[19] because this pump-pump-probe scheme is sampling a different starting geometry/location and energy of the excited state potential.

The increase in ionization, shown in the upper red trace of Figure 26, at short delay time is precisely the reason an IR-UV gain spectrum was possible for this system, in fact the gain spectrum can be collected at any UV wavelength in the range but the longer the UV wavelength the lower the background UV-UV signal and the greater the signal to noise for the IR spectrum. This data is also showing that there is a large geometry change between the ground and excited states and that FC overlap exists between the hot ground state and at least one but possibly multiple excited states. This is an interesting property of thymine and uracil that is not reported previously in the other nucleobases. If the hot ground state created through internal conversion can then be re-excited by UV photons then the role of a hydrogen bonding solvent becomes even more important for removing excess vibrational energy after internal conversion. Experiments have shown that the triplet quantum yield for thymine in acetonitrile is three orders of magnitude larger than in water.[5] Perhaps this could be due to the different solvent not removing vibrational energy from the hot ground state as efficiently thus leading to re-absorption of UV photons after internal conversion and an overall larger triplet population.

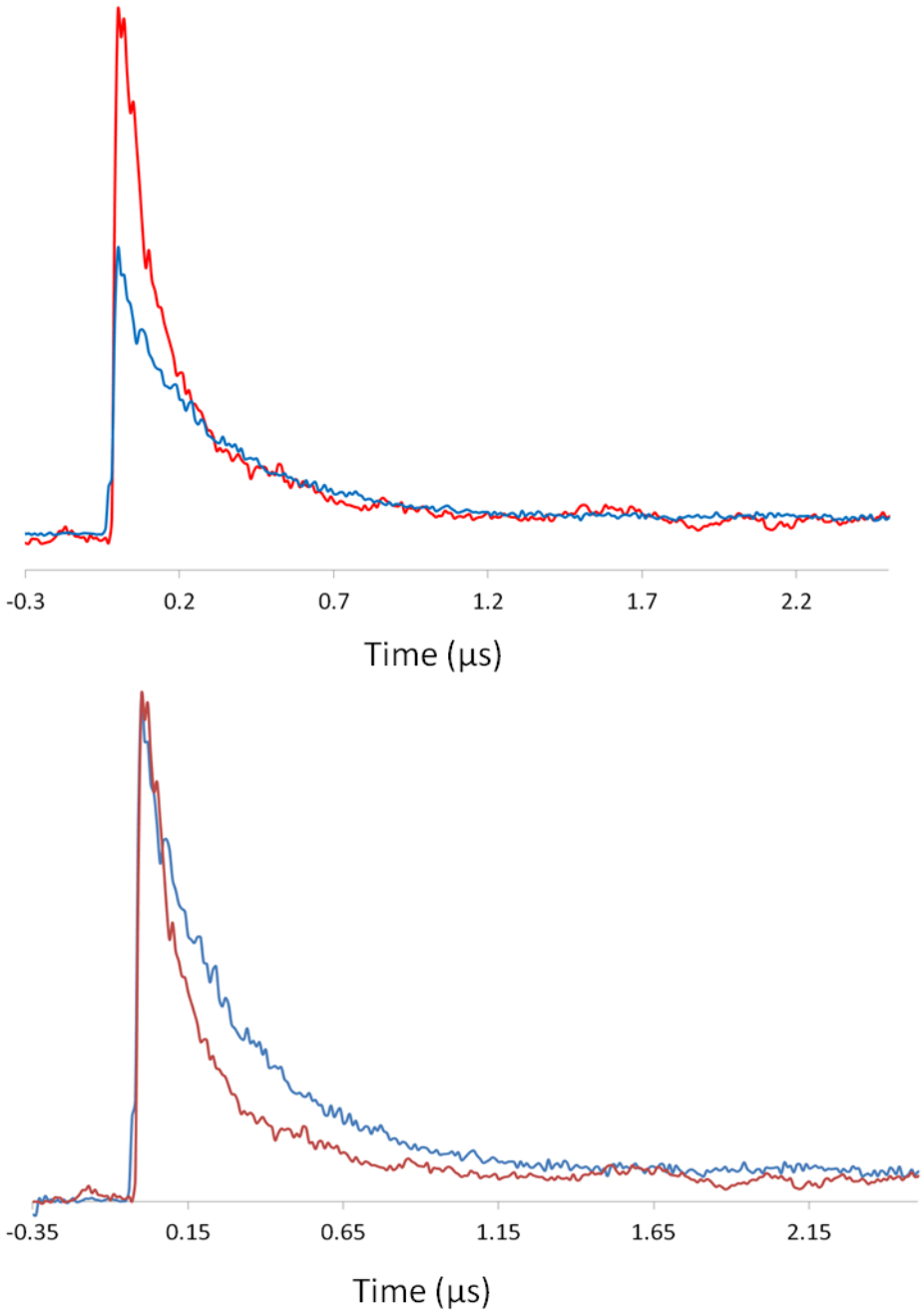


Figure 26. (Top) Unscaled (Bottom) scaled (Red) IR-UV-UV pump-pump-probe lifetime curves (Blue) UV-UV pump-probe lifetime curves at 37021 cm^{-1}.

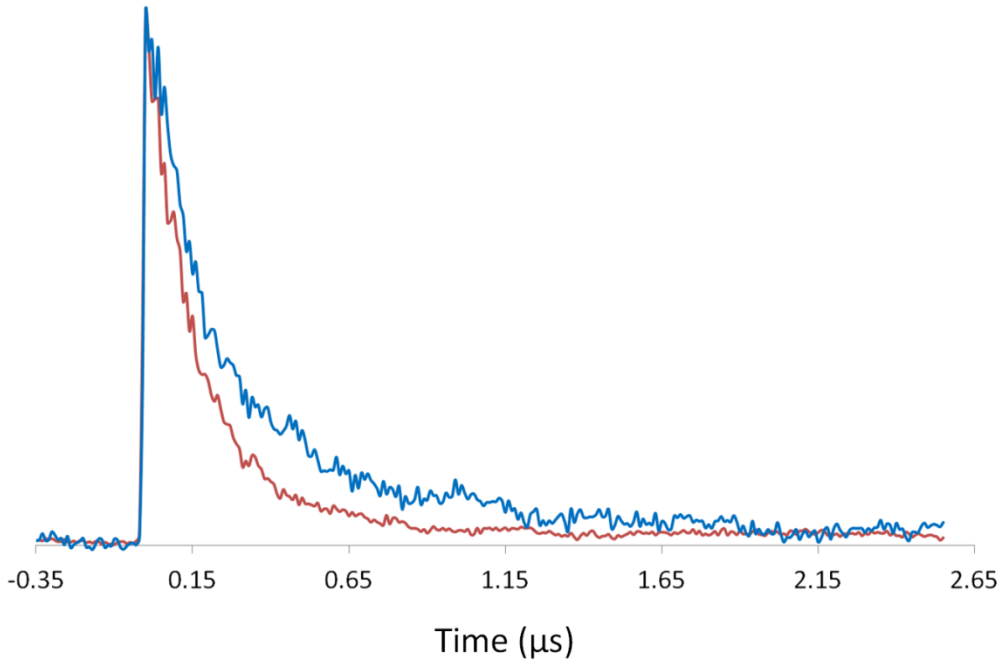


Figure 27. (Red) IR-UV-UV pump-pump-probe lifetime curves (Blue) UV-UV pump-probe lifetime curves at 36364 cm^{-1}.

An experiment was also preformed where the UV laser excites the S_0 $^1(\pi\pi^*)$ state, then 10 ns later the IR laser pumps the excited state vibrationally followed by a variable delay ionization laser. Here again a shorter excited state lifetime is measured. No gain in the microsecond long state could be measured, so population transfer from the nanosecond to microsecond component is not clear. Although if the microsecond component is the hot ground state then the ionization efficiency is very low and a small gain in the population would not be seen at the low overall ionization measured at >1.5 μs.

6. Comparison to Previous Experiments in the Frequency and Time Domain

As mentioned previously the UV spectra for thymine and uracil was first collected nearly 20 years ago.[17] The spectra presented here are qualitatively the same. Both exhibit

broad spectra with some discernible features. Conclusions as to the source of the broad spectrum in the previous publication appear to be quite accurate. They concluded that a large geometry change and multiple electronic states close in energy could explain the broad spectrum. Since then both of these properties have been predicted in nearly all theoretical calculations.[39,40,12,11] In this experiment we have collected data that support the previous findings but it appears that the Franck-Condon overlap contributing to the broad spectrum may not be from the zero point energy ground state but from the vibrationally hot ground state. This is supported by the analysis in section 7 and by a TD-DFT calculation attempted by us which failed to predict any Franck-Condon overlap between the ground state and the S_2 ($^1\pi\pi^*$) state.

More recently the IR spectrum for 1-methylthymine along with lifetime measurements for thymine and 1-methylthymine were reported by Kunitski et al.[19] Just as in our experiment they were unable to obtain an IR-UV hole burning spectrum for thymine. Although it was possible for 1-methylthymine in the ground and excited states and they were able to show that only the keto tautomer was present. As well they reported lifetime measurements for both molecules in the femtosecond and nanosecond time range with fixed UV excitation wavelengths of 267 nm and 266 nm respectively. The lifetime measured in our experiment at the highest energy, 37012 cm^{-1} (270 nm), is in agreement with the previous finding of 280 ± 30 ns. Although we see a long lived component with a > 8 μs lifetime that was not reported previously, this difference could arise from the choice of probe lasers as they used 800 nm photons from a femtosecond amplifier. With the excited state IR spectra they concluded that the long lived state for 1-methylthymine is triplet not singlet in character by comparing the excited state spectrum to theory.[6] The DFT calculation

used predicts a small red shift of 1 cm^{-1} for N3-H in the T_1 ($^3\pi\pi^*$) and a blue shift of 36 cm^{-1} for N3-H in the S_1 ($^1n\pi^*$). Their experiment showed a 10 cm^{-1} red shift making the assignment for 1-methylthymine clearer than for thymine.

A vibronic spectrum of 5-methyl-2-hydroxy pyrimidine collected by Lobsiger et al.[41] shows similar spectral characteristics to thymine and uracil. In that spectrum an in-plane vibronic progression which at higher energies leads to an ISC shows how the vibronic transitions can be resolved until the vibrational density of states becomes too large. For that case excited state mixing occurs at multiple geometries close to the ground state geometry showing sharp spectral features that at higher energies lead to a broad absorption, ~ 3600 cm^{-1} above the origin. IR-UV double resonance spectra of the ground and excited states show that only the enol tautomer is contributing to the spectrum. An IR spectrum of the excited state taken at 50 ns after UV excitation shows a slight broadening and is red shifted by 24 cm^{-1} relative to the ground state but the assignment of the spectrum for the S_1 ($^1n\pi^*$) or T_1 ($^3\pi\pi^*$) is not made from a frequency analysis. Instead the assignment of the long lived T_1 ($^3\pi\pi^*$) state is made by measuring the ionization signal at multiple probe wavelengths and delay times for a fixed pump wavelength. These pump-probe measurements show that at short delay times and longer probe wavelengths the initially excited S_1 ($^1n\pi^*$) state decays in a few ns to a state lived for > 5 μs and is ~ 2998 cm^{-1} below the initially excited state. The vibronic bands are assigned by both RI-CC2 and B3LYP harmonic frequencies to the S_1 ($^1n\pi^*$) excited state and the >5 μs lived state is assigned to the T_1 ($^3\pi\pi^*$). The IR spectrum collected in that experiment was at a delay of 50 ns after UV excitation and was thus the T_1 ($^3\pi\pi^*$) state. The UV and IR spectrum of 5-methyl-2-hydroxy pyrimidine is of interest because of the similar way in which vibronic features at low UV energies becomes a broad

continuum as excitation becomes high enough in energy to couple the bright S_1 ($^1n\pi^*$) state to the T_1 ($^3\pi\pi^*$). As well they were able to assign vibrational frequencies of the S_1 to the vibronic spectrum and these modes had similar movement to structural changes between the ground state and excited S_1 ($^1n\pi^*$) state. This is similar to what is shown for thymine and uracil in section 7 except here the ground state frequencies are assigned.

7. Comparison of Previous Calculations to the Vibronic Spectrum

Calculations for thymine and uracil have been completed at many levels of theory and by multiple groups. They range from estimating minimum energy structures of ground and excited states to internal conversion and intersystem crossing geometries/energies as well as excited state dynamics simulations. Two reviews of have been recently published on this topic for the photo-physics of nucleobases specifically. One covering the theory behind the calculations[42] and another analyzing the differences and similarities of the methods used over the last 15 years.[43]

As mentioned in the results and discussion section 1. a TD-DFT Franck-Condon analysis was performed by us to further evaluate the features in the UV spectrum of both thymine and uracil. Due to a very large geometry change no Franck-Condon overlap was predicted yet there are obvious features in the high resolution spectra. A comparison of the ground state harmonic frequencies using the B97D functional and 6-31G* basis set have allowed for the assignment of 14 features in thymine and 11 features for uracil. Only certain frequencies and combinations or overtones of those frequencies are seen in the spectrum. If we are assuming that the peaks within the broad spectrum are Franck-Condon allowed transitions from a hot ground state to low lying vibronic states then such transitions could come about from linear multi-photon events over the 10 ns perturbation time. A process such as UV

absorption, subsequent ultrafast IC and re-absorption from the hot ground state could occur. If this were the case we would expect that the UV action spectrum would include vibronic excitations leading to the long lived state in which ionization occurs and that vibronic contributions which lead back to the ground state through ultrafast IC would be absent. With this idea in mind the next two sections compare calculations of excited state minimum energy structures and conical intersections/intersystem crossings structures with the nuclear motions of the frequencies assigned to the vibronic features. Pitfalls of this analysis and other possible scenarios leading to the features in the UV spectrum will be discussed at the end of the analysis.

All calculations of the S_2, S_1, and T_1 minima for thymine and uracil are characterized by in-plane contraction of the C4-C5 bond, and lengthening of the C4=O and C5-C6 to different extents depending on the method used and the state studied. There are some disagreements as to the planarity of the S_2 and S_1 states but the T_1 is significantly non planer due to pyramidalization at C6. As well a methyl rotation is predicted between the S_0 and S_2/T_1 states from two different computational methods.[12,6] The calculated energies for both molecules is in agreement in all high level calculations with S_2 being the highest followed by S_1 then T_1.

7.1 Thymine

A well known conical intersection (CI) for thymine involves pyramidalization of the C5 atom with an out of plane bend of the C5-CH3.[9,14] Assignment of a dip feature in the broad UV spectrum at 36945.4 cm^{-1} to the vibrational mode with similar nuclear motion (740.3 cm^{-1}) gives a relative starting location for the assignment of other peaks in the spectrum to vibrational frequencies. The assignments are shown in Figure 28.

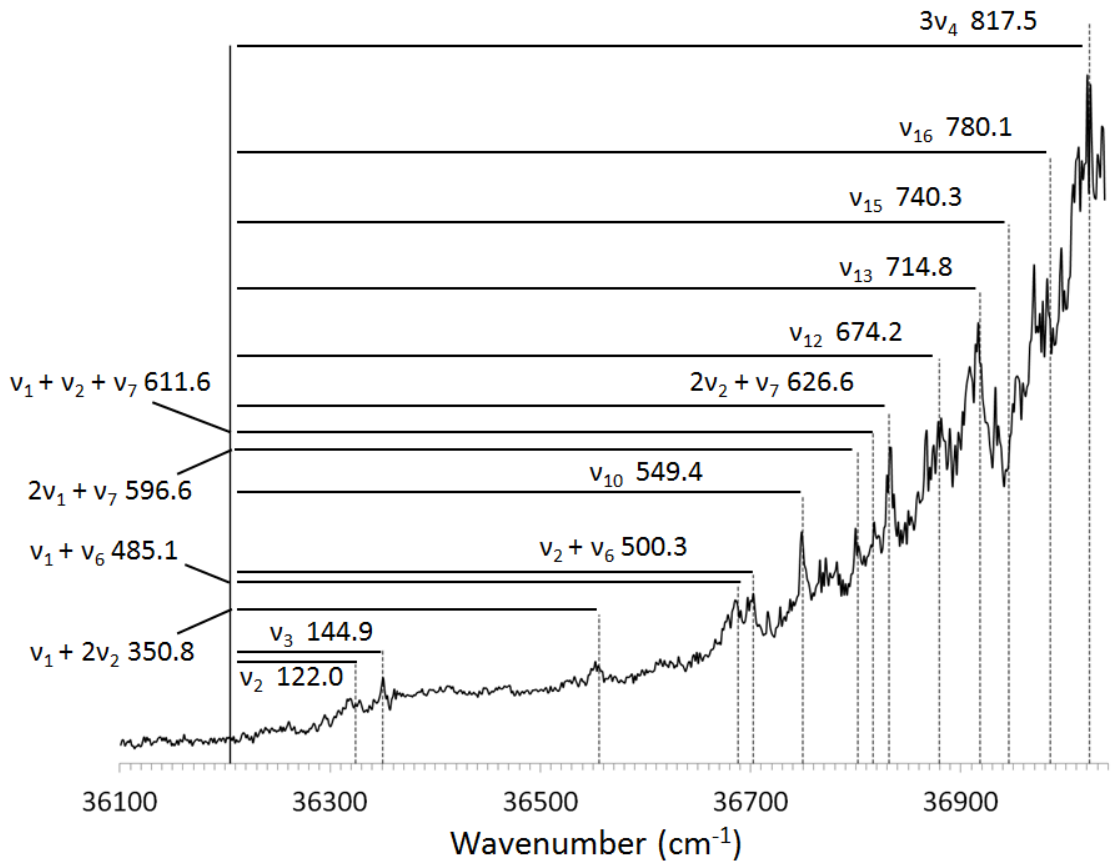


Figure 28. UV spectrum for thymine showing the ground state vibrational mode assignments to the vibronic features. The harmonic frequencies were calculated by DFT using the B97D functional and 6-31G* basis set.

The fundamental bands of ν_2, ν_3, ν_{10}, ν_{12}, ν_{13} and ν_{15} are identified and all except one are out of plane ring deformations and out of plane bends. The exception is ν_{13} which corresponds to an in-plane ring mode with most of the movement being a stretch of the C4-C5 bond (714.8 cm^{-1}). All calculations of the S_1 ($^1n\pi^*$) minimum energy structure at three different levels of theory have shortened C4-C5 bonds by ~ .1 angstrom in comparison to the ground state structures.[44,14,6] This would suggest that the nuclear motion in the hot ground

state creates FC overlap allowing absorption to the S_1 ($^1n\pi^*$) directly or the vibronic transition to the S_2 ($^1\pi\pi^*$) state potential surface activates internal conversion to the S_1 ($^1n\pi^*$) state. If the long lived state is T_1 ($^3\pi\pi^*$) then this would suggest that the triplet state would be populated via the S_1 ($^1n\pi^*$) after internal conversion from the S_2 ($^1\pi\pi^*$). Pathways for ISC from S_1 ($^1\pi\pi^*$) to T_1 ($^3\pi\pi^*$) have been identified by various computational methods,[7,6] of these pathways the S_2 ($^1\pi\pi^*$) to S_1 ($^1n\pi^*$) to T_1 ($^3\pi\pi^*$) was concluded to be dominant by Marian and coworkers.[6]

The three lowest energy modes assigned, ν_1, ν_2 and ν_3, exhibit small out of plane ring deformation accompanied by methyl rotation. The modes ν_2 and ν_3 are identified as the two lowest energy peaks in the spectrum and ν_1 is only seen in combination with ν_6 and ν_7. Combinations of ν_2 with ν_6 and ν_7 also match peaks in the spectrum. Again calculations by Marian and coworkers show that the S_1 ($^1n\pi^*$) minimum is a double minimum where both states are planner but one has a 60 degree rotation of the methyl group with respect to the ground state. Other calculations by Blancafort and coworkers,[45] and Lischka and coworkers predict a methyl rotation on the relaxation from the FC region to the S_2 ($^1\pi\pi^*$) minimum,[12] no rotation of the S_1 ($^1n\pi^*$) minimum but a similar rotation of the T_1 ($^1\pi\pi^*$) minimum as for the S_2. In the former work one of the two CIs identified was an indirect path from the S_2 ($^1\pi\pi^*$) F-C point to the S_1 ($^1n\pi^*$) state which involves methyl rotation as it crosses the S_2 ($^1\pi\pi^*$) minimum and the S_2/S_1 CI. Six of the twelve vibronic transitions identified in the UV spectrum involve either directly or through a combination band a methyl rotation from the three lowest energy modes. Just as the other frequency assignment of the in-plane stretch of C4-C5 this would implicate S_1 as an intermediate state and a possible vibrational motion that would not only activate the S_2/S_1 crossing but the S_1/T_1 as well.

The frequencies for the N2-H and N4-H bends labeled ν_{10} and ν_{12} in Figure 24 also show small contributions to the vibronic spectrum. Although no calculations are directly implicating these modes the T_1 ($^3\pi\pi^*$) state is predicted to be significantly non planner where as S_1 ($^1n\pi^*$) is planner. This could be either some small population that undergoes direct ISC from the S_2 ($^1\pi\pi^*$) or through other unidentified modes in the intermediate S_1 ($^1n\pi^*$). The overtone band of $2\nu_4$ is nearly isoenergetic with ν_{10} and is an in plane methyl bend which is also a likely assignment to this peak.

Good agreement with the ground state harmonic frequencies and the features in the broad UV spectrum indicate that the electronic excitation from the ground state could be occurring from the hot ground state to the FC allowed S_2 ($^1\pi\pi^*$) potential surface. This type of absorption combined with trapping in the dark state could explain the broad spectrum measured by nanosecond multi-photon ionization. Although the presence of the broad continuous absorption in combination with the sharper features could be a combination of absorption from the hot and cold ground states to different excited states as will be discussed after the uracil analysis.

7.2 Uracil

All high level calculations concerning uracil and thymine are in agreement that the lowest lying excited singlet and triplet states are similar in their structure, dynamics and relative energies. Thus we would expect to see a similar UV spectrum and indeed we do. The same type of ground state harmonic frequency comparison made for thymine also fits for uracil. Starting with the frequency ν_{13} (785.4 cm^{-1}), which involves symmetric C5-H and C6-H out of plane bending motions and assigning it to the dip feature at 37079.4 cm^{-1} allows

for the assignment of many other vibrational modes to features in the spectrum (Figure 29). This mode is assigned to the lower energy dip feature due to its similarities with the S_2/S_0 conical intersection. As well another dip feature 139.9 cm^{-1} to the blue is assigned to ν_{14} the asymmetric C5-H and C6-H bending mode. The ultrafast IC pathway of S_2/S_0 is similar in every calculation and involves pyramidalization of C6 with a C5-H out of plane bend also accompanied by an out of pane distortion. Using the energy of these two modes and subtracting their frequencies from the UV spectrum gives the origin of the spectrum. A total of 11 features can be assigned to various fundamental, combination and overtone bands. As can be seen in the Figure 29 there is a peak at the origin for uracil that was not seen in thymine.

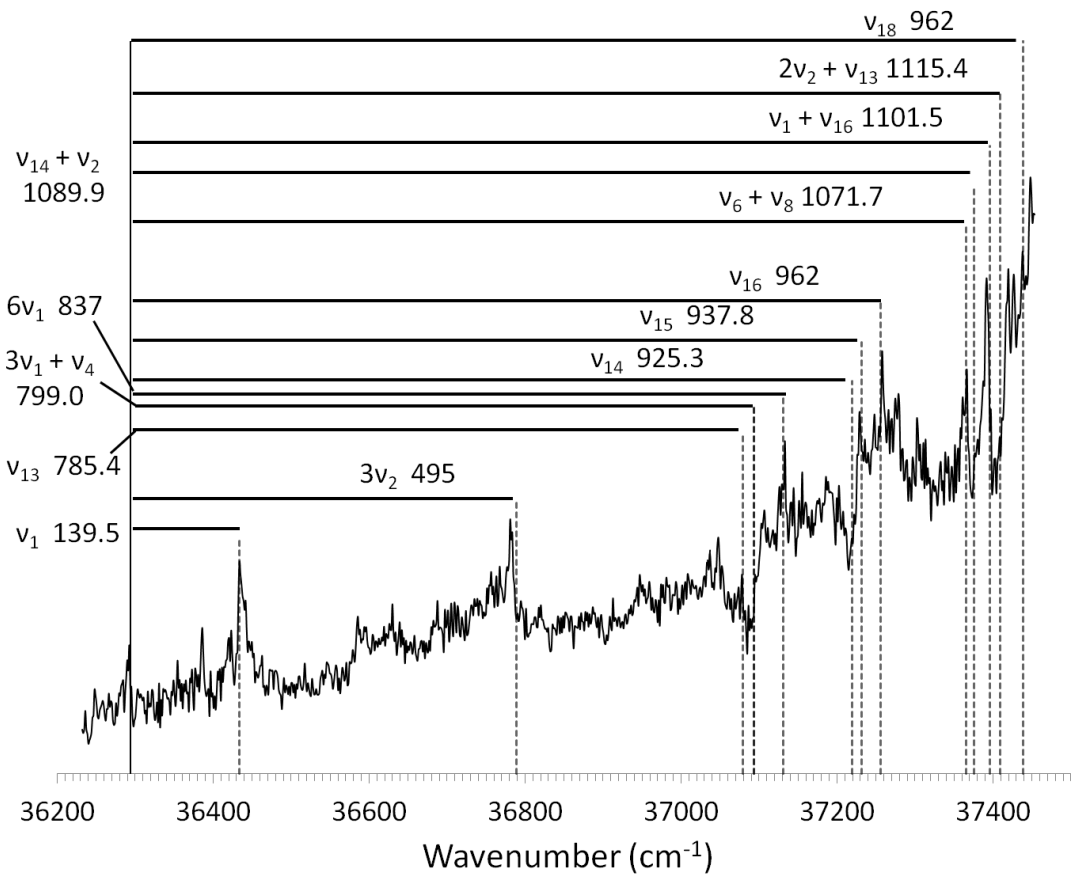


Figure 29. UV spectrum for uracil showing the ground state vibrational mode assignments to the vibronic features. The harmonic frequencies were calculated by DFT using the B97D functional and 6-31G* basis set.

The next peak corresponds to ν_1, which is a symmetric out of plane ring bending motion with the largest motion in the N4 nitrogen. In a calculation by Matiska[46] an S_2/S_1 CI was located where the structure has a symmetric out of plane bend that is nearly identical to the turning point of ν_1. This frequency is identified as the lowest energy vibronic transition and appears to be involved in combination and overtone bands with three other peak assignments. Just as for thymine this would suggest that the S_1 state is a possible

intermediate between the S_2 ($^1\pi\pi^*$) state and the T_1 ($^3\pi\pi^*$) or perhaps the S_1 is borrowing intensity from the S_2 and the absorption is direct into the S_1.

The other three fundamental modes that can be assigned are ν_{15}, ν_{16} and ν_{18}. The lower energy mode at 937.8 cm^{-1} is an in plane ring stretch of C4-C5. This mode is the same as for thymine which was assigned to the most discernible peak in the spectrum. Slightly higher in energy is the ν_{16} (962.0 cm^{-1}) fundamental mode, it is an in plane deformation of the N1, N3 and C5 atoms. This mode stretches the C4=O from ~1.2 Angstroms to ~1.38 Angstroms during its cycle which again would be a characteristic of the changes between the S_0 and S_2, S_1, T_1 states. Although at this turning point of motion, the rest of the bond lengths and angles are distorted from equilibrium and would not be expected to give good FC overlap with the S_1 or T_1 minimum directly but with no knowledge of the excited state frequencies no conclusions can be made about why this mode contributes to the spectrum. It should be noted that the 7th overtone of ν_1 would be 10 cm^{-1} the blue of this assignment and still well within the broad peak for the ν_{16} assignment.

The rest of the features are assigned to combination and overtone frequencies. In the lower energy region a peak is seen just to the red of the 3rd overtone for ν_2 which is an asymmetric out of plane ring distortion mostly involving pyramidalization of the N1 and C5 atoms. The peak being just to the red of the third overtone would make sense for harmonic frequency calculated overtones. No direct comparison can be made for this mode but the S_2 ($^1\pi\pi^*$) and T_1 ($^3\pi\pi^*$) and in some calculations S_1 ($^1n\pi^*$) are predicted to be non planer and as for ν_{16} this mode is activating N1 and C5 out of plane distortions. It should also be

mentioned here that two S_2/S_1 CIs and two S_1/S_0 CIs identified by Nachtigallova et. al.[11] display asymmetric out of plane distortions in uracil.

Another combination band involving the 3rd overtone for ν_1 coupled to ν_4 and lies 13.6 cm^{-1} above ν_{13} and could be contributing to the large dip feature. The ν_1 mode, as described above is a out of plane pyramidalization of the N3 which creates a boat like structure near its turning point similar to a S_2/S_1 CI.[46] The ν_4 mode is also a symmetric type ring deformation but has large pyramidalization at C6 which would be similar to the ultrafast S_2/S_0 CI, as well many calculations predict a similar S_1/S_0 CI for pyramidalization at the C5 and C6 atoms.[11] Because this coupled mode predicts movement along possibly separate pathways it is difficult to estimate how this vibronic excitation decays but either way its energy is the same as the lower energy dip feature and may prove to be useful when deciding starting locations for dynamics calculations on the excited potential energy hyper surface (PEH).

Because of the lower overall low ion signal of uracil *vs* thymine and the lowest energy vibrational motions being generally similar to both S_2/S_0 and S_1/S_0 CIs the spectrum above ~ 37100 cm^{-1} becomes difficult to interpret. Other than the fundamental modes already discussed many overtone and combination bands seem to show what would be expected for the overall signal but individual assignments are not as clear. It seems that the ν_1 and ν_2 modes couple with other modes and depending on the coupling this either leads to direct internal conversion to the ground state or the dark state. The high energy region of the UV spectrum in Figure 30 illustrates this conclusion.

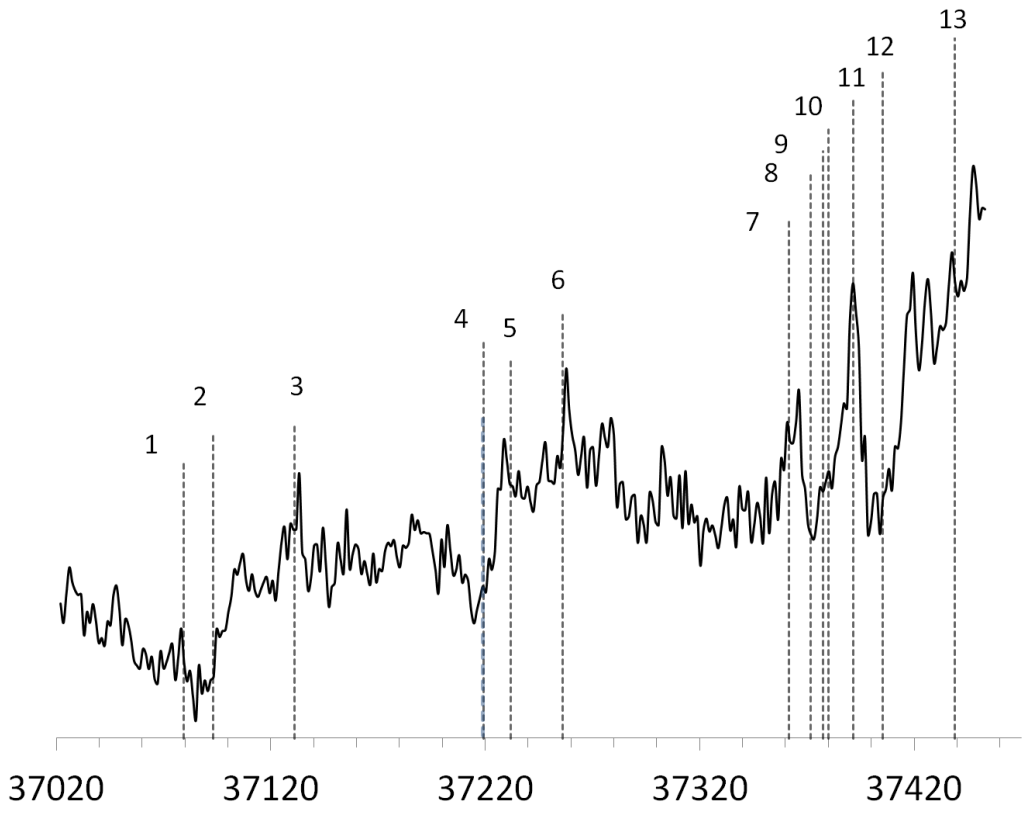


Figure 30. High energy region of the UV spectrum for uracil. The peak assignments are sequential to increasing energy (cm^{-1}) to the following vibrational modes. (1) ν_{13} 785.4 (2) $3\nu_1 + \nu_4$ 799.0 (3) $6\nu_1$ 837 (4) ν_{14} 925.3 and $\nu_1 + \nu_{13}$ 924.9 (5) ν_{15} 937.8 (6) ν_{16} 937.8 (7)* $\nu_6 + \nu_8$ 1071.7 (8)* $\nu_3 + \nu_{10}$ 1081.9 (9)* $\nu_4 + \nu_{10}$ 1087.6 (10)* $\nu_2 + \nu_{14}$ 1090.3 (11)* $\nu_1 + \nu_{16}$ 1101.5 (12)* $2\nu_2 + \nu_{13}$ 1115.4 (13) 1145.1. The frequencies with asterisks are scaled by -4 cm^{-1} in the figure.

Just above 37360 cm^{-1} a cluster of six assignments (7-12) are made around two dip features with a large peak between them. To the low energy side a peak is assigned to $\nu_6 + \nu_8$ (7) neither of these modes are assigned alone but ν_6 is an in plane deformation that bends

the carbonyls, stretches the C4=O while at the same time stretching the C5-C6 bond and ν_8 is an N-H out of plane bend with large oscillator strength. The next assignment to the low energy side in the first dip is a combination of $\nu_3 + \nu_{10}$ (8). The ν_{10} mode is the C5-H out of plane bend accompanied by pyramidalization at C5 while ν_3 is an in plane symmetric bending motion of the carbonyls. Also within this dip is $\nu_4 + \nu_{10}$ (9) The combination of these two modes essentially creates the S_2/S_0 CI structure predicted in nearly every computational study. The next assignment only 2.3 cm^{-1} higher in energy and also within the dip feature can be assigned to two sets of combination modes, $\nu_{14} + \nu_2$ (10) and $\nu_{13} + \nu_1 + \nu_2$ (10), that are 0.4 cm^{-1} apart. These are both combinations of the modes assigned to the other dips in the spectrum and the calculations predict that CIs of S_2/S_0 and S_1/S_0 include a bend of C5 and out of plane distortions so combinations of ν_{14} and ν_{13}, symmetric and antisymmetric C5 and C6 bends, with ν_1 and ν_2, symmetric and antisymmetric out of plane ring distortions could be directly connecting the S_2/S_0 through this type of vibronic transition. The peak at the center of the dips corresponds to $\nu_{16} + \nu_1$ (11) where ν_{16} is the in plane ring mode previously assigned and described above. ν_1 being the symmetric out of plane bend. The last assignment for the second dip is $\nu_{13} + 2\nu_2$ (12) again like the first dip this would be C5-H and C6-H bends accompanied by asymmetric ring distortions from ν_2. As can be seen by the assignments in this region the number of possible combination and overtone bands matching with possible features makes a definitive conclusion about any single mode not possible without calculations. But the trend overall for uracil seems to be that there are more features assigned to modes that would create similar structures to the S_2/S_0 ultrafast IC than for thymine. This combined with the two lowest energy modes which

appear to be coupling with either the modes leading to the dark state or the hot ground state seems to show a lower probability for populating the triplet state when comparing uracil to thymine. Which is supported experimentally by the ionization on the nanosecond time scale being 10 times lower for uracil than for thymine.

Summary

The ground state and excited state spectra for thymine and uracil show the biologically relevant diketo tautomer to be the only species present in a laser desorption seeded molecular beam. The anharmonic frequency calculations along with the experimentally measured red shift of frequencies for vibrationally warm molecules would strongly suggest that the long lived states for isolated thymine and uracil are indeed triplet in character.

The resolved features within the broad spectra for both molecules are reported here for the first time. The assignment of many features in the UV spectrum to ground state harmonic frequencies suggest that absorption from S_0 ($^1\pi\pi$) to S_2 ($^1\pi\pi^*$) occurs from the hot ground state rather than the cold ground state. This could occur by a three photon process of UV absorption, subsequent ultrafast IC to the ground state and then re-absorption from the hot ground state to a favorable FC point in the excited state which then undergoes either internal conversion back down to the ground state or proceeds to the dark state and is ionized by the third UV (193 nm) photon. Comparison of the UV spectra to harmonic frequencies along with past experiments and calculations suggests that the dark T_1 ($^3\pi\pi^*$) state is populated via the intermediate S_1 $^1(n\pi^*)$ state. Other possible scenarios for absorption include the mixing of excited states that leads to a borrowing of intensity of the S_1 $^1(n\pi^*)$ vibronic transitions from the S_2 $^1(\pi\pi^*)$ bright state. The sharp features over the broad background signal can be explained by this borrowing of intensity where the broad

spectrum is hot ground state absorption into either S_1 or S_2 and the sharp features are cold ground state vibronic transitions into S_1 $^1(n\pi^*)$. Although a comparison of the S_1 vibrational frequencies of thymine from a previous calculation[6] (DFT B3LYP/TZVP) does not match the vibronic features neither does the ground state frequencies by the same method where as the ground state DFT B97D/6-31G* does match so S_1 excited state frequencies using B97D functional may prove that the S_1 is responsible for the vibronic features arising from the cold ground state. A recent calculation by Krylov and coworkers[47] suggests that the systematic over estimation of vertical excitation energies for uracil by many methods is due to vibronic excitation directly into S_1 rather than S_2 which may be what our experimental data is showing. Overall this type of absorbance in combination with hot ground state absorbance to the S_2 $^1(\pi\pi^*)$ state is perhaps the best explanation of the experimental data.

The absence of a ground state hole burning spectrum and the gain in ionization after IR excitation of the ground state shows that both thymine and uracil have large changes in geometry between the ground and excited state, which confirms the large structural changes predicted in calculations. As well this shows that there is FC overlap between the vibrationally excited ground state and the S_2 ($^1\pi\pi^*$) or other unknown excited state, thus giving experimental evidence for the validity of the ground state harmonic frequency analysis. The differences seen in the overall number of frequencies corresponding to peaks and dips in the UV spectrum for thymine and uracil suggest that thymine has more low energy modes that create FC overlap with the absorbing excited state and more of those vibronic transitions lead to the T_1 ($^3\pi\pi^*$) state. For thymine this appears to be due to a rotation of the methyl group between the S_0 ($^1\pi\pi$) and S_2 ($^1\pi\pi^*$) minimum energy structures and the three lowest energy modes that undergo nuclear motion of this type. Many

combination bands are formed with these modes that lead to the T_1 ($^3\pi\pi^*$) state. Uracil also shows extensive combination and overtone bands from its low energy vibrational modes but coupling of these modes with two or possibly three modes that lead to direct IC, rather than just one for thymine, gives uracil better protection against triplet formation that is inherent in its vibrational motions.

References

1. Beukers, R.; Berends, W., ISOLATION AND IDENTIFICATION OF THE IRRADIATION PRODUCT OF THYMINE. *Biochimica Et Biophysica Acta* **1960,** *41* (3), 550-551.
2. Lamola, A. A.; Mittal, J. P., SOLUTION PHOTOCHEMISTRY OF THYMINE AND URACIL. *Science* **1966,** *154* (3756), 1560-&.
3. Lamola, A. A.; Yamane, T., SENSITIZED PHOTODIMERIZATION OF THYMINE IN DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1967,** *58* (2), 443-&.
4. Hare, P. M.; Middleton, C. T.; Mertel, K. I.; Herbert, J. M.; Kohler, B., Time-resolved infrared spectroscopy of the lowest triplet state of thymine and thymidine. *Chem. Phys.* **2008,** *347* (1-3), 383-392.
5. Salet, C.; Bensasson, R., STUDIES ON THYMINE AND URACIL TRIPLET EXCITED-STATE IN ACETONITRILE AND WATER. *Photochemistry and Photobiology* **1975,** *22* (6), 231-235.
6. Etinski, M.; Fleig, T.; Marian, C. A., Intersystem Crossing and Characterization of Dark States in the Pyrimidine Nucleobases Uracil, Thymine, and 1-Methylthymine. *J. Phys. Chem. A* **2009,** *113* (43), 11809-11816.
7. Serrano-Perez, J. J.; Gonzalez-Luque, R.; Merchan, M.; Serrano-Andres, L., On the intrinsic population of the lowest triplet state of thymine. *J. Phys. Chem. B* **2007,** *111* (41), 11880-11883.
8. Climent, T.; Gonzalez-Luque, R.; Merchan, M.; Serrano-Andres, L., On the intrinsic population of the lowest triplet state of uracil. *Chem. Phys. Lett.* **2007,** *441* (4-6), 327-331.
9. Merchan, M.; Gonzalez-Luque, R.; Climent, T.; Serrano-Andres, L.; Rodriuguez, E.; Reguero, M.; Pelaez, D., Unified model for the ultrafast decay of pyrimidine nucleobases. *J. Phys. Chem. B* **2006,** *110* (51), 26471-26476.
10. Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallova, D.; Hobza, P.; Lischka, H., Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (50), 21453-21458.
11. Nachtigallova, D.; Aquino, A. J. A.; Szymczak, J. J.; Barbatti, M.; Hobza, P.; Lischka, H., Nonadiabatic Dynamics of Uracil: Population Split among Different Decay Mechanisms. *J. Phys. Chem. A* **2011,** *115* (21), 5247-5255.
12. Szymczak, J. J.; Barbatti, M.; Hoo, J. T. S.; Adkins, J. A.; Windus, T. L.; Nachtigallova, D.; Lischka, H., Photodynamics Simulations of Thymine: Relaxation into the First Excited Singlet State. *J. Phys. Chem. A* **2009,** *113* (45), 12686-12693.
13. Lamola, A. A., EXCITED STATE PRECURSORS OF THYMINE PHOTODIMERS. *Photochemistry and Photobiology* **1968,** *7* (6), 619-&.
14. Zechmann, G.; Barbatti, M., Photophysics and deactivation pathways of thymine. *J. Phys. Chem. A* **2008,** *112* (36), 8273-8279.
15. Zhao, H. M.; Liu, K. H.; Song, D.; Su, H. M., Physical Quenching in Competition with the Formation of Cyclobutane Pyrimidine Dimers in DNA Photolesion. *J. Phys. Chem. A* **2014,** *118* (39), 9105-9112.
16. Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernandez, C. E.; Kohler, B., DNA Excited-State Dynamics: From Single Bases to the Double Helix.

In *Annual Review of Physical Chemistry*, Annual Reviews: Palo Alto, 2009; Vol. 60, pp 217-239.

17. Brady, B. B.; Peteanu, L. A.; Levy, D. H., THE ELECTRONIC-SPECTRA OF THE PYRIMIDINE-BASES URACIL AND THYMINE IN A SUPERSONIC MOLECULAR-BEAM. *Chem. Phys. Lett.* **1988,** *147* (6), 538-543.
18. Busker, M.; Nispel, M.; Haber, T.; Kleinermanns, K.; Etinski, M.; Fleig, T., Electronic and vibrational spectroscopy of 1-methylthymine and its water clusters: The dark state survives hydration. *ChemPhysChem* **2008,** *9* (11), 1570-1577.
19. Kunitski, M.; Nosenko, Y.; Brutschy, B., On the Nature of the Long-Lived "Dark" State of Isolated 1-Methylthymine. *ChemPhysChem* **2011,** *12* (10), 2024-2030.
20. Chernyshova, I. V.; Kontros, J. E.; Markush, P. P.; Shpenik, O. B., Excitation of lowest electronic states of the uracil molecule by slow electrons. *Opt. Spectrosc.* **2012,** *113* (1), 5-8.
21. Chernyshova, I. V.; Kontros, E.; Markush, P. P.; Shpenik, O. B., Excitation of lowest electronic states of thymine by slow electrons. *Opt. Spectrosc.* **2013,** *115* (5), 645-650.
22. He, Y. G.; Wu, C. Y.; Kong, W., Decay pathways of thymine and methyl-substituted uracil and thymine in the gas phase. *J. Phys. Chem. A* **2003,** *107* (26), 5145-5148.
23. Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A., Electronic relaxation dynamics in DNA and RNA bases studied by time-resolved photoelectron spectroscopy. *Phys. Chem. Chem. Phys.* **2004,** *6* (10), 2796-2801.
24. Honig, R. E.; Woolston, J. R., LASER-INDUCED EMISSION OF ELECTRONS, IONS, AND NEUTRAL ATOMS FROM SOLID SURFACES. *Appl. Phys. Lett.* **1963,** *2* (7), 138-139.
25. Karas, M.; Bachmann, D.; Hillenkamp, F., INFLUENCE OF THE WAVELENGTH IN HIGH-IRRADIANCE ULTRAVIOLET-LASER DESORPTION MASS-SPECTROMETRY OF ORGANIC-MOLECULES. *Anal. Chem.* **1985,** *57* (14), 2935-2939.
26. Karas, M.; Gluckmann, M.; Schafer, J., Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J. Mass Spectrom.* **2000,** *35* (1), 1-12.
27. Meijer, G.; Devries, M. S.; Hunziker, H. E.; Wendt, H. R., LASER DESORPTION JET-COOLING OF ORGANIC-MOLECULES - COOLING CHARACTERISTICS AND DETECTION SENSITIVITY. *Applied Physics B-Photophysics and Laser Chemistry* **1990,** *51* (6), 395-403.
28. Kantrowitz, A.; Grey, J., A HIGH INTENSITY SOURCE FOR THE MOLECULAR BEAM .1. THEORETICAL. *Rev. Sci. Instrum.* **1951,** *22* (5), 328-332.
29. Smalley, R. E.; Wharton, L.; Levy, D. H., MOLECULAR OPTICAL SPECTROSCOPY WITH SUPERSONIC BEAMS AND JETS. *Accounts Chem. Res.* **1977,** *10* (4), 139-145.
30. C. Moller, M. P., Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934,** *46* (7), 618-622.
31. Hattig, C. W., *Adv. Quant. Chem.* **2005,** *50*, 37-50.
32. Kohn, A.; Hattig, C., Analytic gradients for excited states in the coupled-cluster model CC2 employing the resolution-of-the-identity approximation. *J. Chem. Phys.* **2003,** *119* (10), 5021-5036.

33. Hattig, C., Geometry optimizations with the coupled-cluster model CC2 using the resolution-of-the-identity approximation. *J. Chem. Phys.* **2003,** *118* (17), 7751-7761.
34. Dunning, T. H., GAUSSIAN-BASIS SETS FOR USE IN CORRELATED MOLECULAR CALCULATIONS .1. THE ATOMS BORON THROUGH NEON AND HYDROGEN. *J. Chem. Phys.* **1989,** *90* (2), 1007-1023.
35. Barone, V., Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.* **2005,** *122* (1).
36. Cooley, J. W., An Improved Eigenvalue Corrector Formula for Solving the Schrödinger Equation for Central Fields. *Mathematics of Computation* **1961,** *15*, 363.
37. Lobsiger, S.; T'Rachsel, M. A.; Frey, H. M.; Leutwyler, S., Excited-State Structure and Dynamics of Keto-Amino Cytosine: The (1)pi pi* State Is Nonplanar and Its Radiation less Decay Is Not Ultrafast. *J. Phys. Chem. B* **2013,** *117* (20), 6106-6115.
38. Makarov, A. A.; Petrova, I. Y.; Ryabov, E. A.; Letokhov, V. S., Statistical inhomogeneous broadening of infrared and Raman transitions in highly vibrationally excited XY6 molecules. *J. Phys. Chem. A* **1998,** *102* (9), 1438-1449.
39. Lorentzon, J.; Fulscher, M. P.; Roos, B. O., THEORETICAL-STUDY OF THE ELECTRONIC-SPECTRA OF URACIL AND THYMINE. *J. Am. Chem. Soc.* **1995,** *117* (36), 9265-9273.
40. Shukla, M. K.; Mishra, P. C., A gas phase ab initio excited state geometry optimization study of thymine, cytosine and uracil. *Chem. Phys.* **1999,** *240* (3), 319-329.
41. Lobsiger, S.; Frey, H. M.; Leutwyler, S., Supersonic jet UV spectrum and nonradiative processes of the thymine analogue 5-methyl-2-hydroxypyrimidine. *Phys. Chem. Chem. Phys.* **2010,** *12* (19), 5032-5040.
42. Matsika, S.; Krause, P., Nonadiabatic Events and Conical Intersections. In *Annual Review of Physical Chemistry, Vol 62*, Leone, S. R.; Cremer, P. S.; Groves, J. T.; Johnson, M. A., Eds. Annual Reviews: Palo Alto, 2011; Vol. 62, pp 621-643.
43. Angelo Giussani, J. S.-M., Daniel Roca-Sanjuan and Manuela Merchan, Excitation of Nucleobases from a Computational Perspective I: Reaction Paths. Springer: Verlag Berlin Heidelberg, 2013.
44. Perun, S.; Sobolewski, A. L.; Domcke, W., Conical intersections in thymine. *J. Phys. Chem. A* **2006,** *110* (49), 13238-13244.
45. Asturiol, D.; Lasorne, B.; Robb, M. A.; Blancafort, L., Photophysics of the pi,pi* and n,pi* States of Thymine: MS-CASPT2 Minimum-Energy Paths and CASSCF on-the-Fly Dynamics. *J. Phys. Chem. A* **2009,** *113* (38), 10211-10218.
46. Matsika, S., Radiationless decay of excited states of uracil through conical intersections. *J. Phys. Chem. A* **2004,** *108* (37), 7584-7590.
47. Epifanovsky, E.; Kowalski, K.; Fan, P. D.; Valiev, M.; Matsika, S.; Krylov, A. I., On the electronically excited states of uracil. *J. Phys. Chem. A* **2008,** *112* (40), 9983-9992.

IV. New Method for the Direct Analysis of Xanthine Stimulants in Archaeological Vessels

Introduction

In the vast arsenal of analytical chemical techniques that are routinely used in archaeometry, the majority is particularly suited to elemental analysis. Tools such as neutron activation,[1] energy dispersive x-ray (EDX),[2] atmospheric pressure chemical ionization (APCI),[3] and accelerator mass spectrometry[4] all make major contributions to identification and characterization of archaeological materials. On the other hand, analysis of organic compounds in archeological samples is more challenging, generally requiring combinations of techniques and often yielding less specific identifications. One such combination of techniques which is widely used is hybrid instrumentation that pairs chromatographic separation with mass spectrometric detection. For example, GC/MS often identifies general types of compounds, such as food groups, with great success and interpretive power. While these techniques constitute a useful approach, they generally lack the capability to identify unique compounds. This shortcoming particularly applies to complex organic compounds and mixtures. Therefore other techniques geared towards identifying specific molecular markers have the potential to provide valuable complementary information. Here we describe a new technique for organic tracer molecule analysis for archaeometry, in a specialized form of laser mass spectrometry. This technique combines the selectivity of resonant laser spectroscopy with the sensitivity of mass spectrometry and is therefore simultaneously highly specific and sensitive.[5] In the following sections we will detail the

technique, followed by a first example of application to the identification of xanthine markers in beverage residues in Maya and Mississippian pottery sherds.

Two-step Laser Mass Spectrometry

The instrument is schematically shown in Figure 31. A more detailed description is given in chapter II and reference six.[6]

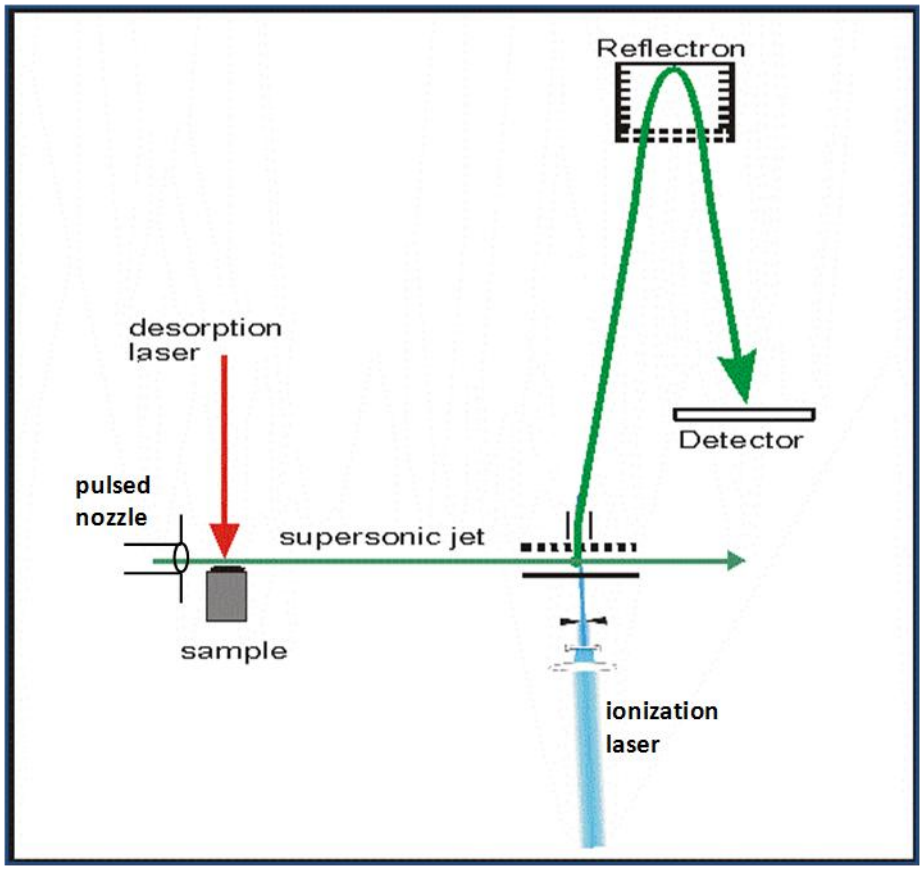


Figure 31. Schematic of the molecular beam mass spectrometer**.**

The first step is laser desorption which makes it possible to vaporize large and complex or thermally labile molecules intact. The sample can be either in the form extract deposited on a sample substrate, or it can be actual pottery material from which we laser desorb

directly. We combine this step with jet cooling by entraining the desorbed molecules in a supersonic expansion of an inert gas.[7] Samples are placed on a translating stage inside the high vacuum chamber directly in front of a pulsed molecular beam valve. The jet entrainment causes a very efficient cooling of the internal degrees of freedom of the molecules. The cooling makes it possible to perform high resolution spectroscopy while at the same time stabilizing the molecule, permitting its detection at the parent molecular mass. We use a number of different wavelengths (266 nm, 532 nm, 1064 nm) to desorb from a wide variety of substrates, including metals, graphite and the sample material itself. We typically use desorption laser fluence of the order of 1.0 mJ/cm^2 in 10 ns laser pulses.[8] The desorption laser is focused using either a cylindrical lens or a spherical lens, generating desorption spot sizes of 0.50 mm x 3.0 mm and 0.75 mm diameter respectively. The next step is ionization, for which we employ photoionization in several ways: resonance enhanced two-photon ionization (R2PI)[9] with one or two colors and single photon ionization. The subsequent ions are detected by a reflectron time of flight mass spectrometer.

R2PI

By using resonance enhanced two-photon ionization (R2PI), we are able to combine optical spectroscopy with mass spectrometry. This dramatically enhances the specificity for selected compounds and allows for distinction of structural isomers, tautomers and enantiomers.[10] We excite the cold molecule to the first electronic excited state by a photon from a tunable dye laser. The excited molecule is subsequently ionized by another photon, either from the same dye laser (1 color - 2 photon ionization), or from another laser (2 color - 2 photon ionization). The first and resonant photon comes from the doubled output of a

lumonics HD-300 tunable dye laser with maximum spectral line width 0.003 nm and pulse energy 0.3-0.7 mJ. For two color ionization the second photon comes from another dye laser, which in these experiments was fixed at 308 nm. (1.5-2 mJ/pulse). The ions are detected in a reflectron time-of-flight mass spectrometer. By scanning the wavelength of the first photon, while monitoring a specific ion mass, we obtain a mass selected excitation spectrum. Once the spectroscopy of a given compound is known, we can tune the wavelength to a particular molecular resonance for selective ionization of the individual analyte. This makes it possible to detect compounds within a matrix of many other compounds and to do so at very low concentrations. Resonant ionization not only selects for a specific compound, it can also select for specific isomers. We have demonstrated that with this technique we can detect compounds at the femtomol level and in favorable cases even down to the 100 attomol level.[5b] We can further improve R2PI sensitivity by two color ionization. Typically the absorption cross sections for the first and second step are 10^{-17} cm^2 and 10^{-19} cm^2, respectively. To maintain optimum selectivity it is undesirable to significantly saturate the first step, forcing us to use a laser fluence that is too low by two orders of magnitude to saturate the second step. Therefore we can improve overall sensitivity without sacrificing selectivity if we employ a different wavelength for the second step at higher laser fluence. We have demonstrated this principle for perylene, obtaining an overall 0.25 photoionization efficiency resulting in a 30 femtogram detection limit.[5b]

The combination of laser spectroscopy and mass spectrometry provides analytical information in two dimensions: wavelength and mass. Generally spectroscopic resolution, which is typically fractions of wave numbers is several orders of magnitude higher than the mass resolution that can be obtained in conventional mass spectrometry.[10] To fully capitalize

on these advantages it is necessary that the spectroscopy of the analyte molecule is known. We can find a needle in a haystack, provided we know what the needle looks like.

Application: Cacao Traces in Maya Pottery

The xanthine alkaloids are a class of central nervous system stimulants that are consumed daily worldwide. They have a rich history of recorded use in many civilizations. The plants that contain them are found on every continent except Antarctica. In the Americas there are several types of *Ilex* commonly known as Holly, *Theobroma Cacao* in Central and South America, *Camilla Sinensis* or tea in Asia and *Coffea Arabica* in the Middle East and Africa. Today coffee is the second highest traded commodity on the world market beat only by crude oil. In prerecorded history stimulant containing plants likely played an important role in any society that had knowledge of these plants. Specific methyl xanthines can serve as tracer compounds for these types of beverages in ancient pottery. Figure 3 shows the three methylxanthine derivatives monitored in this work: 3,7-dimethylxanthine or theobromine, 1,3-dimethylxanthine or theophylline, and 1,3,7-dimethylxanthine or caffeine. Their relative abundances vary for different plant types, with theobromine prevalent in cacao but less abundant in Holly.

Previous analysis of pottery from Central America has identified theobromine in residue from the inside of ceramic vessels from Honduras, Guatemala, and Belize dating from 1500 BCE to 480 CE.[3,11,12] The identifications were made by gas chromatography/mass spectrometry (GC/MS) and high performance liquid chromatography/mass spectrometry (HPLC/MS). In each of these cases theobromine was the tracer molecule identified thus the identification of a cacao beverage was facile considering the only other theobromine containing plant on the continent is Ilex cassine which grows 1500-2000 miles away. Crown

et al. identified theobromine in sherds from the American Midwest[13] and in vessels from what is now modern day Southern Illinois, they have also identified all three stimulants in sherds dating as far back as 1050 CE.[14] The authors concluded that *Illex Vomitoria* was the likely source due to historical record of its use in the area but theobromine was found to be the dominant constituent in some of the vessels and the use of *Illex Cassine* could not be ruled out.

Samples and Standards

We studied samples from two different archaeological sites. (1) Thirteen base sherds of unique Late Classic period (c. 600-900 CE) Maya cylinder vessels from the El Pilar area.[1] (2) 7 sherds from the unique early Mississippian period (c. 1100-1200 CE) cylindrical vessels from the Central Illinois River Valley in Fulton County Illinois.[2]

Standards of theobromine, theophylline and caffeine were purchased from Sigma-Aldrich and used without further purification. Standards are directly applied to graphite sample bars as a thin solid layer. The spectra for standards were collected using separate graphite bars to ensure each spectrum is free of any other standards.

Pottery samples were analyzed both from extract and directly from their surfaces. The extracts are made by using a 3:1 mixture of acetone and water. Approximately 400 mg of ground pottery is added to 5 ml of solution and allowed to sit at room temperature for 72 hours. The supernatant liquid is then filtered by a Whatman 13 mm GD/X disposable filter, polypropylene filter media with polypropylene housing, 0.45 mm pore size. The extract is

[1] Provided by Dr. Anabel Ford from the Department of Anthropology, University of California Santa Barbara. We thank the Belize Institute of Archaeology for their support of the El Pilar research.

[2] Provided by Dr. Greg Wilson from the Department of Anthropology, University of California Santa Barbara.

then concentrated by gentle heating (25-35 °C) and dry nitrogen flowing over the surface of the vial. The extracts are concentrated approximately 5 fold then deposited drop wise on disposable stainless steel pegs mounted to the sample bar. Gentle heating (35-45 °C) is used to speed up evaporation of solvent. The dried, concentrated extracts are then immediately inserted into the instrument for analysis. Direct desorption analysis of the samples is done by applying small amounts of either ground or surface pieces to double sided tape mounted on gold sample bars, as shown in Appendix III section 4. A new, clean disposable razor blade was used to scrap sherds fragments directly onto the tape for each sherd tested.

We measured the high resolution UV spectrum of all three molecular markers, allowing us to choose individual wavelengths for optimal spectroscopic identification and selective ionization of the individual isomers, theobromine *vs.* theophylline.[15] Figure 32 shows the different spectra. To signal average we typically collected 15 mass spectra for each wavelength step and averaged 15 wavelength scans for the pottery sherds and 8 wavelength scans for the standards.

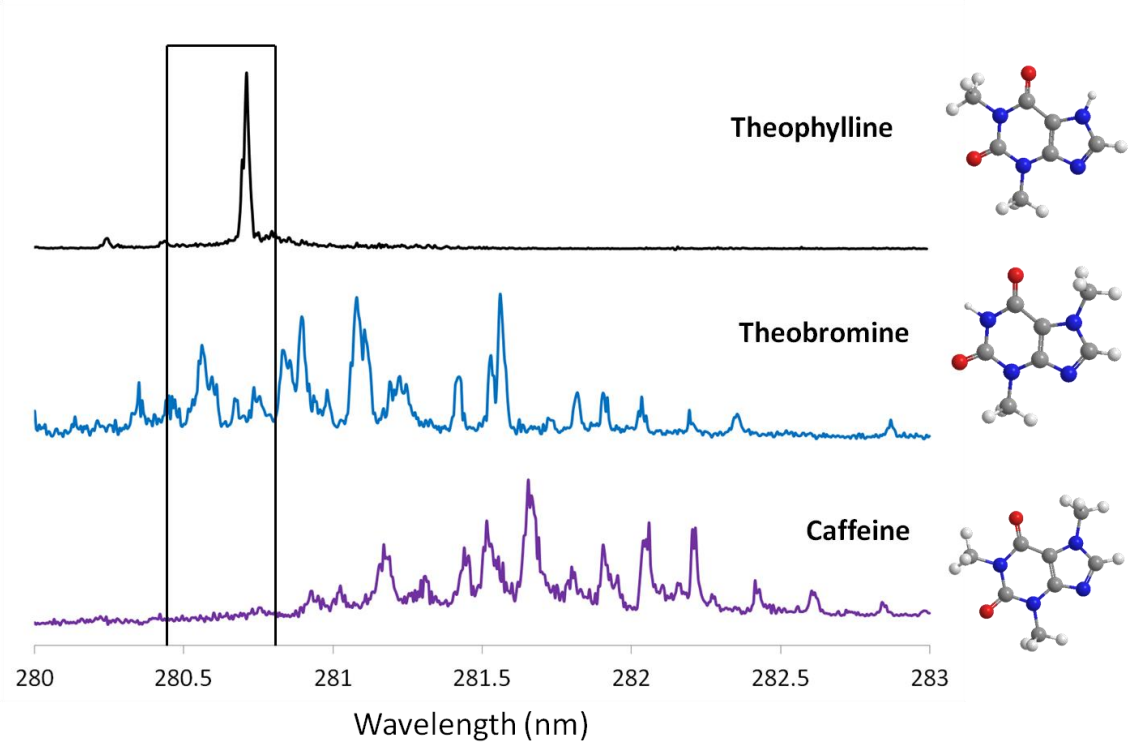


Figure 32. UV spectrum for methylxanthine standards: Theophylline (top, black), Theobromine (middle, blue) and Caffeine (bottom, purple). The back box marks the region scanned for the pottery sherds.

1. Maya Sherds

We obtained positive identification of theobromine in 12 of 13 sherds analyzed from extracts at a resonance wavelength of 281.55 nm. We identified caffeine in only a few extracts at this wavelength. As well theophylline could be identified in a few extracts at the resonant wavelength of 280.69 nm. Using direct desorption, the same 12 of the 13 sherds tested positive for theophylline and caffeine at 280.69 nm. Three vessel sherds were chosen to be tested for theobromine at 281.55 nm by direct desorption all of which tested positive for both theobromine and caffeine. Data showing positive identification using direct

desorption is available in Appendix III. Figure 33 shows a short wavelength scan over the range of 280.45 nm to 280.80 nm which covers resonances for both theophylline and theobromine. Two cylinder vessels, #1 and #4, as well as the pedestal base vessel were chosen for the wavelength scan. Figure 33 shows the spectra of the standards theobromine and theophylline, for clarity only the spectrum of cylinder vessel #4 is shown but both cylinder vessels spectra are similar.

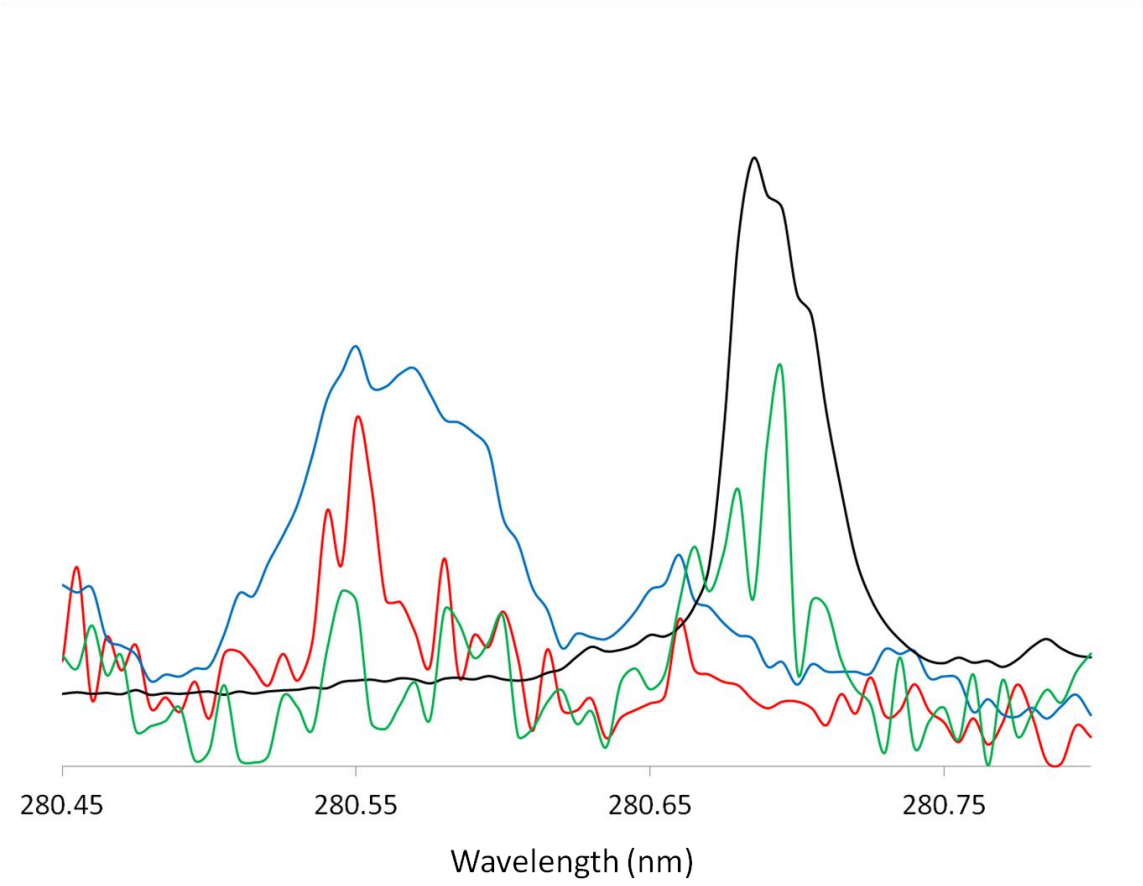


Figure 33. UV spectrum for standards; Theophylline (black), Theobromine (blue) and two Maya vessels; a cylindrical vessel #4 (green) and pedestal base vessel #7 (red). The spectra for the archaeological vessels is scaled ×20 relative the standards spectra.

The spectrum for the pedestal base vessel is distinctly different from that of the cylinder vessels. Its spectrum tracks the standard for theobromine while the cylinder vessel's spectrum tracks mostly the standard for theophylline. A small peak for theobromine is also evident in the cylinder vessels at 280.55 nm. Over the same wavelength range the mass channel for caffeine was monitored (Figures 34 and 35). The pedestal base vessel shows nearly the same signal of caffeine to theobromine/theophylline stimulants whereas the cylinder vessels show nearly ten times the signal of caffeine to theobromine/theophylline. In both figures the y-axis scale is the same. This result suggests that the pedestal base vessel could have been used in a different manner then the cylinder vessels.

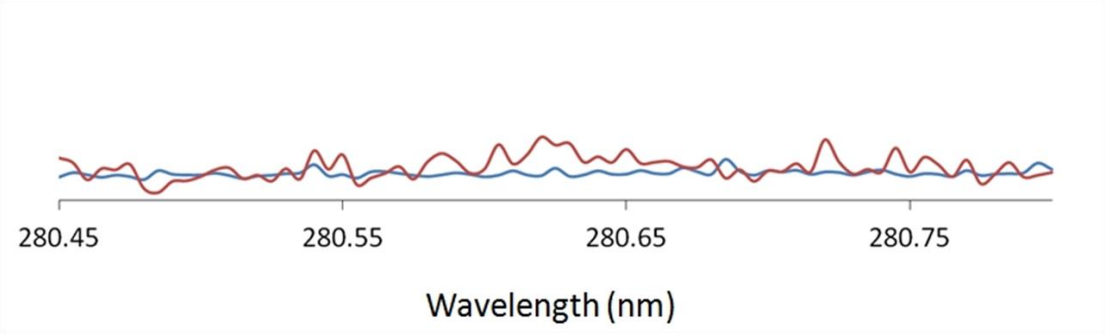


Figure 34. UV spectrum of Maya pedestal base vessel. (Red) Caffeine mass channel. (Blue) Theobromine/theophylline mass channel.

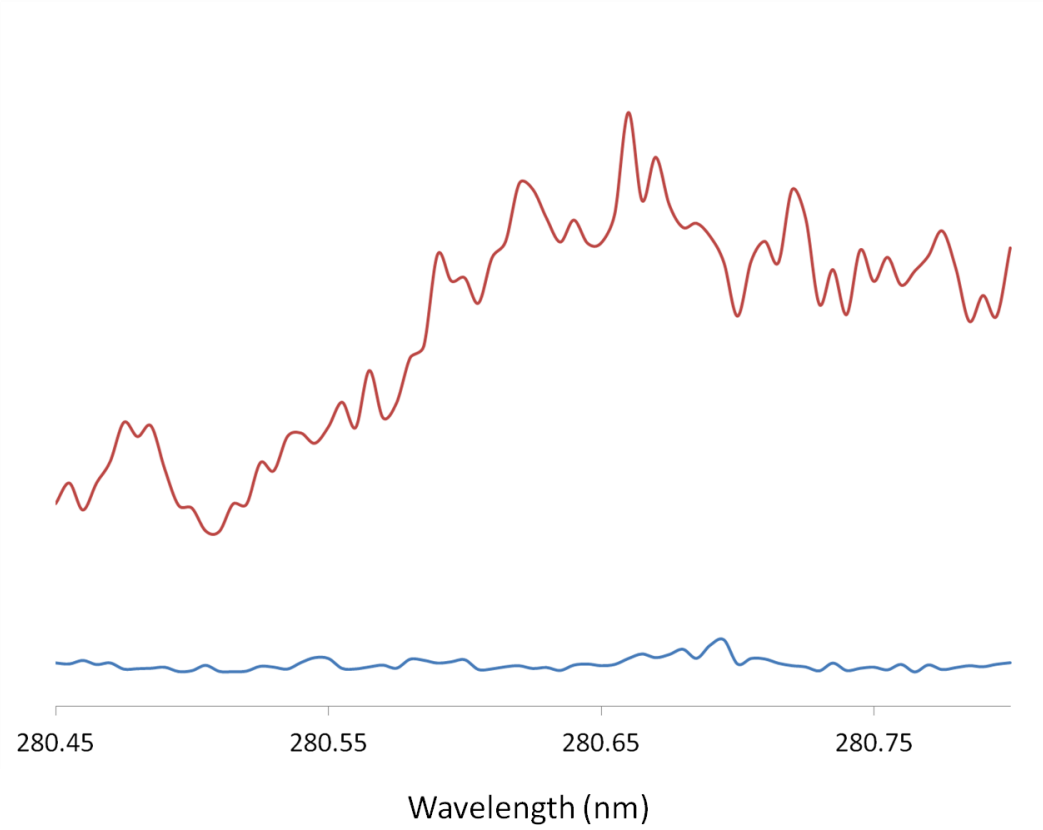


Figure 35. UV spectrum of Maya cylinder vessel. (Red) Caffeine mass channel. (Blue) Theobromine/theophylline mass channel.

Other possible factors that may affect the observed differences in relative abundance of the methyl xanthines include the following: The two types of sherds exhibited visibly different types of clay (Figure 36) and so there is the possibility of one type of clay preferentially binding one stimulant over the other or the desorption laser preferentially volatilizing specific derivatives from the different matrices. For the analysis of extracts there is the additional complication of the risk of differential extraction. The other published technique for analysis of cacao in pottery sherds uses extraction with various solvents followed by LC/MS[3,13,12]. In these experiments theobromine is the only xanthine identified.

Figure 36. (Top) Pedestal base vessel #7. (Bottom) cylinder vessels.

In our attempts to optimize the extraction of theobromine from the sherds a mixture of acetone and water was found to give the best results. With these extractions we were able to identify theobromine in 12 of the 13 vessels and caffeine in only a few of them. The larger relative amount of caffeine identified in the common vessels by direct laser desorption vs. extracted samples suggests modification of the ratios during extraction. Possible reasons include either the extraction not removing all the caffeine from the sample or caffeine degrading or volatilizing during concentration of the extract. Caffeine is known to have the

highest solubility in water of the three xanthines despite having one fewer hydrogen bonding site.[16] Therefore different relative amounts of caffeine could also reflect a different history of exposure to moisture, different histories of use of the vessels and different preparations of the beverages.

2. Mississippi Sherds

We analyzed 7 sherds from the Central Illinois River Valley by the direct laser desorption/resonant laser ionization technique. All 7 samples showed strong positive results for Caffeine at the two resonant wavelengths 280.69 nm (theophylline) and 281.55 nm (theobromine). 6 of the 7 samples tested positive for theophylline and 6 of the 7 tested positive for theobromine at their respective resonant wavelengths. These results are in qualitative agreement with Crown et al.[14] Although this method is not yet quantitative to better than within an order of magnitude the laser desorption technique is showing more caffeine relative to theobromine/theophylline. Data confirming these identifications are available in Appendix III section 3.

Summary

We presented a new method for analysis of molecular markers for stimulant containing beverages. Several pottery sherds from both Central and North America tested positive for all three xanthine alkaloids; caffeine, theobromine and theophylline. Although the method is not yet quantitative, relative amounts of caffeine to theobromine or caffeine to theophylline are consistent throughout the data set at their respective resonant wavelengths for the direct desorption method. The Maya sherds tested positive for theobromine by extraction and direct desorption analysis. Direct desorption revealed larger amounts of caffeine than

extraction. This method is more sensitive than previous methods by identifying all three xanthine stimulants in multiple samples with much smaller sample sizes. The two major advantages of this technique include; reduced sample sizes needed for positive identification and the ability to directly analyze complex matrices such as clay from pottery without extraction of the molecular markers prior to analysis.

Future projects will include the analysis of sherds from ancient Nubia. Here the class of tracer molecules know as polyphenols will be the target analytes. As well mock pottery samples made from clays of their respective archaeological sites will be spiked with standards of tracer molecules. These standard samples will be tested for determining detection limits, quantification limits and the possibility of differential recovery of the analytes using the direct desorption method.

References

1. Glascock, M. D.; Neff, H., Neutron activation analysis and provenance research in archaeology. *Meas. Sci. Technol.* **2003,** *14* (9), 1516-1526.
2. Bronk, H.; Rohrs, S.; Bjeoumikhov, A.; Langhoff, N.; Schmalz, J.; Wedell, R.; Gorny, H. E.; Herold, A.; Waldschlager, U., ArtTAX - a new mobile spectrometer for energy-dispersive micro X-ray fluorescence spectrometry on art and archaeological objects. *Fresenius J. Anal. Chem.* **2001,** *371* (3), 307-316.
3. Hurst, W. J.; Tarka, S. M.; Powis, T. G.; Valdez, F.; Hester, T. R., Archaeology: Cacao usage by the earliest Maya civilization. *Nature* **2002,** *418* (6895), 289-290.
4. Piperno, D. R.; Flannery, K. V., The earliest archaeological maize (Zea mays L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98* (4), 2101-2103.
5. (a) Arrowsmith, P.; de Vries, M. S.; Hunziker, H. E.; Wendt, H. R., Pulsed laser desorption near a jet orifice: concentration profiles of entrained perylene vapor. *Appl. Phys. B* **1988,** *46*, 165-73; (b) Meijer, G.; de Vries, M. S.; Hunziker, H. E.; Wendt, H. R., Laser desorption jet-cooling of organic molecules. Cooling characteristics and detection sensitivity. *Appl. Phys. B* **1990,** *51*, 395-403; (c) Nir, E.; Hunziker, H. E.; de Vries, M. S., Fragment Free Mass Spectrometric Analysis with Jet Cooling/VUV Photoionization. *Anal. Chem.* **1999,** *71*, 1674-1678; (d) de Vries, M. S.; Elloway, D. J.; Wendt, H. R.; Hunziker, H. E., Photoionization mass spectrometer with a microscope laser desorption source. *Review of Scientific Instruments.* **1992,** *63*, 3321-3325; (e) Mahajan, T. B.; Plows, F. L.; Gillette, J. S.; Zare, R. N.; Logan, G. A., Comparison of microprobe two-step laser desorption/laser ionization mass spectrometry and gas chromatography/mass spectrometry studies of polycyclic aromatic hydrocarbons in ancient terrestrial rocks. *J. Am. Soc. Mass Spectrom.* **2001,** *12* (9), 989-1001; (f) Gillette, J. S.; Ghosh, U.; Mahajan, T. B.; Zare, R. N.; Luthy, R. G., Microprobe laser mass spectrometry studies of polycyclic aromatic hydrocarbon distributions on harbor sediments and coals. *Isr. J. Chem.* **2001,** *41* (2), 105-110; (g) Mahajan, T. B.; Ghosh, U.; Zare, R. N.; Luthy, R. G., Microscale detection of polychlorinated biphenyls using two-step laser mass spectrometry. *Int. J. Mass Spectrom.* **2001,** *212* (1-3), 41-48.
6. Meijer, G.; Devries, M. S.; Hunziker, H. E.; Wendt, H. R., LASER DESORPTION JET-COOLING OF ORGANIC-MOLECULES - COOLING CHARACTERISTICS AND DETECTION SENSITIVITY. *Applied Physics B-Photophysics and Laser Chemistry* **1990,** *51* (6), 395-403.
7. (a) Li, L.; Lubman, D. M., Pulsed laser desorption method for volatilizing thermally labile molecules for supersonic jet spectroscopy. *Rev. Sci. Instrum.* **1988,** *59*, 557-561; (b) Weyssenhoff, H. V.; Selzle, H. L.; Schlag, E. W., Laser-desorbed large molecules in a supersonic jet. *Zeitschrift fur Naturforschung, Teil A* **1985,** *40a*, 674-6.
8. (a) Meijer, G.; de Vries, M. S.; Hunziker, H. E.; Wendt, H. R., Laser desorption jet-cooling spectroscopy of the benzoic acid monomer. *Journal of Physical Chemistry.* **1990,** *94*, 4394-6; (b) Anex, D. S.; de Vries, M. S.; Knebelkamp, A.; Bargon, J.;

Wendt, H. R.; Hunziker, H. E., Resonance-enhanced two-photon ionization time-of-flight spectroscopy of cold perfluorinated polyethers and their external and internal van der Waals dimers. *International Journal of Mass Spectrometry and Ion Processes.* **1994,** *131*, 319-34; (c) Nir, E.; Grace, L. I.; Brauer, B.; de Vries, M. S., REMPI Spectroscopy of Jet Cooled Guanine. *J. Am. Chem. Soc.* **1999,** *121*, 4896-4897; (d) Cohen, R.; Nir, E.; Grace, L. I.; Brauer, B.; de Vries, M. S., Resonance-Enhanced Multiphoton Ionization Spectroscopy of Dipetides. *J. Phys. Chem. A (USA)* **2000,** *104*, 6351-6355; (e) Nir, E.; Imhof, P.; Kleinermanns, K.; de Vries, M. S., REMPI spectroscopy of laser desorbed guanosines. *J. Am. Chem. Soc.* **2000,** *122* (33), 8091-8092; (f) Nir, E.; Muller, M.; Grace, L. I.; de Vries, M. S., REMPI spectroscopy of cytosine. *Chem. Phys. Lett.* **2002,** *355* (1-2), 59-64.

9. Zandee, L.; Bernstein, R. B., multiphoton ionization of a molecular benzene beam. *Journal-of-Chemical-Physics.* **1979,** *70*, 2574-5.
10. Imasaka, T.; Moore, D. S.; Vo-Dinh, T., Critical assessment: Use of supersonic jet spectrometry for complex mixture analysis - (IUPAC technical report). *Pure Appl. Chem.* **2003,** *75* (7), 975-998.
11. Hurst, W. J.; Martin, R. A.; Tarka, S. M.; Hall, G. D., AUTHENTICATION OF COCOA IN MAYA VESSELS USING HIGH-PERFORMANCE LIQUID-CHROMATOGRAPHIC TECHNIQUES. *Journal of Chromatography* **1989,** *466*, 279-289.
12. Henderson, J. S.; Joyce, R. A.; Hall, G. R.; Hurst, W. J.; McGovern, P. E., Chemical and archaeological evidence for the earliest cacao beverages. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (48), 18937-18940.
13. Crown, P. L.; Hurst, W. J., Evidence of cacao use in the Prehispanic American Southwest. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (7), 2110-2113.
14. Crown, P. L.; Emerson, T. E.; Gu, J. Y.; Hurst, W. J.; Pauketat, T. R.; Ward, T., Ritual Black Drink consumption at Cahokia. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109* (35), 13944-13949.
15. Callahan, M. P.; Gengeliczki, Z.; Svadlenak, N.; Valdes, H.; Hobza, P.; de Vries, M. S., Non-standard base pairing and stacked structures in methyl xanthine clusters. *Phys. Chem. Chem. Phys.* **2008,** *10* (19), 2819-2826.
16. Jouyban, A., *Handbook of Solubility Data for Pharmaceuticals*. CRC Press: Boca Raton, FL 33487-2742, 2009.

Appendix I

1. Tables of calculated frequency assignments to experimental spectra

dAMP

Experiment		Calculation
Intensity	Frequency (cm^{-1})	
0.43818	1625.90978	(1617) NH2 Scissor mode
0.06364	1587.52373	(1575) C6-C5 stretch coupled to NH2 scissor
0.06411	1461.10784	(1450) C2-H in plane wag
0.0616	1410.09888	
0.54601	1288.70971	(1248) CH2 bend on sugar coupled with asymmetric O-P-O stretch
		(1240) asymmetric O-P-O stretch coupled with CH2 bend on sugar
0.25162	1248.91759	
0.17947	1215.1607	(1198) C2-H in plane wag
0.46971	1103.86187	(1056) P-O-H bend (free)
		(1117) C-O stretch on hydroxyl of sugar (H bonded to phosphate)
0.3073	1051.73405	(987) C-O on sugar ring couple to P-O (H bonded to 3'OH)
		(977) C-O on sugar ring couple to P-O (H bonded to 3'OH)
0.03698	980.55784	
0.03767	945.39782	
0.43172	826.96101	(719) large P-OH stretch (free)
0.14536	726.45068	(619) large P-OC stretch

dGMP

Experiment		Calculation
Intensity	Frequency (cm-1)	
0.27443	1716.7382	(1719) carbonyl stretch
0.06448	1632.65306	(1669) NH2 scissor mode (H bonded to phosphate)
0.21087	1574.80315	(1584) NH2 scissor mode coupled to ring movement (all in plane)
0.03058	1324.50331	(1336) N5-H inplane bend
0.03183	1312.33596	(1312) inplane ring mode coupled to C1-H and N5-H in plane bends
0.15889	1249.21924	(1194) asymmetric O-P-O stretch
0.20895	1095.89041	(1030) C5'-O stretch coupled to PO-H bend
		(1068) large in plane ring mode coupled to P-O-H bend
0.07	1050	(992) asymmetric O-P-O coupled with C5'-O bonding sugar and phosphate
0.04608	1020.40816	(997) C-O stretch on sugar
0.08814	823.04527	(768) Large P-OH stretch (free)

dCMP

Experiment		Calculation
Intensity	Frequency (cm^{-1})	
0.30041	1673.64017	(1675) Carbonyl stretch
0.34799	1626.01626	(1627) NH2 in plane scissor
0.23194	1520.91255	(1522) C5-C6 stretch
0.12622	1454.54545	(1465) C4-N3 stretch
0.10414	1384.08304	
0.231	1324.5	(1337) C4-NH2 stretch
0.62388	1277.95527	(1229) asymmetric O-P-O stretch
		(1250) CH2 wag on sugar coupled to asymmetric O-P-O stretch
		(1289) C6-N1 stretch
0.519	1075.26882	(1012) symmetric O-P-O stretch (H bonded to O3'H of sugar)
0.39873	1036.26943	
0.03455	966.18357	
0.45377	796.81275	(728) P-OH stretch (free)
		(789) O3' bend/ H bonded to phosphate
0.08998	716.33238	(622) P-OC stretch
0.02421	667.55674	

dTMP

Experiment		Calculation
Intenstiy	Frequency (cm^{-1})	
0.52007	1708.56333	(1700) C2=O carbonyl stretch
		(1685) C4=O carbonyl stretch
0.29923	1284.61606	(1244) asymmetric O-P-O stretch coupled with CH2 bends on sugar
		(1246) asymmetric O-P-O stretch coupled with CH2 bends on sugar
		(1225) asymmetric O-P-O stretch coupled with CH2 bends on sugar
0.17142	1092.55255	(1120) C-OH stretch (H bonded 3'OH to phosphate)
		(1067) C-O bond from sugar to phosphate
		(1050) P-O-H bend (free)
		(1011) symmetric O-P-O stretch
0.16349	1049.01319	
0.20598	809.21175	(731) P-OH stretch (free)
0.05186	771.51875	(681) P-OC5' stretch
0.03352	722.3369	(635) N3-H bend
		(616) P-OC5' stretch

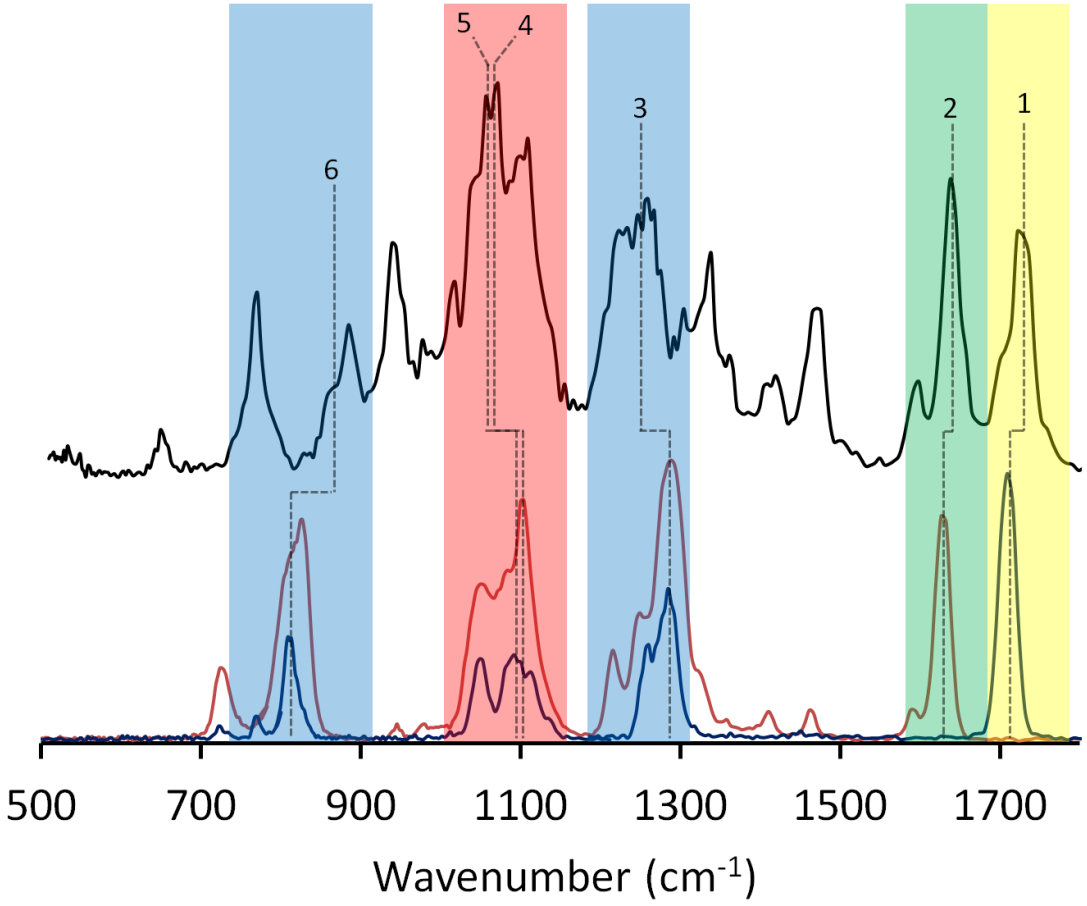
dAMP-dTMP

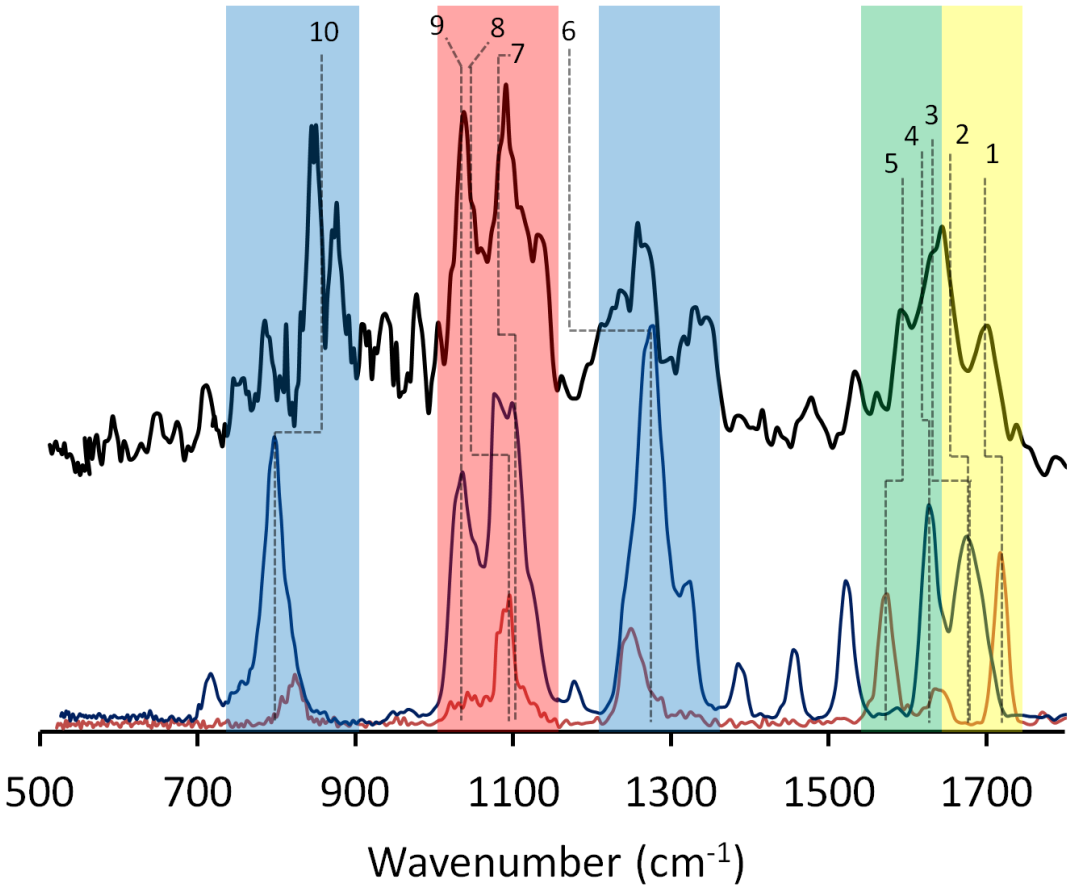
Experiment		Calculation
Intensity	Frequency (cm^{-1})	
0.48733	1722	(1721) Carbonyl on dTMP
0.58833	1638	(1614) NH2 scissor mode
0.19773	1599	(1568) NH2 scissor mode coupled with C5-C6 stretch on dAMP
0.08435	1498 (shoulder)	(1477) OH , CH bend both on sugar
		(1470) large ring movement all C-N and C-C
0.33927	1471	(1452) C6-H inplane bend on dAMP
		(1435) CH3 on dTMP
0.20987	1419	(1433) large asym CH3 on dTMP
		(1406) CH bends on sugars
0.24939	1362	(1351,1358) large ring movement on dAMP and CH wag on same sugar
0.44413	1339	(1340) CH in plane bend on Tmp some CH3 bend on same sugar
		(1336) All CH bends on Amp sugar some OH
0.33905	1304	(1304) CH and CH2 on sugar of dAMP
		(1297) large all CNs on dAMP
0.48868	1270-1220	(1228) NH2 in plane wag dAMP, (1220-1270) mostly CH and CH2 bends and wags some dTMP ring movement
		(1209) double OH bend on phosphate of dAMP (double H bonded), P=O stretch
		(1197) C8-H wag on dAMP coupled to CH2 on dAMP sugar
		(1181) Asymmetric O=P-O stretch on dTMP (double H bonded)
		(1169) P=O stretch coupled to double OH bend on phosphate of dAMP
0.66531	1110	(1038) C1'-O4' on sugar of Amp coupled with CH2 bends
		(1017) phosphate bound C5'-O stretch on both Amp and Tmp coupled with both sugars
		(1012) P-O stretch on Tmp coupled with P-O-H bends on both phosphates
0.77223	1072.33428	(991) C5'-O coupled with asymmetric P-O on Tmp
0.38997	1018.49464	(985) NH2 in plane wag on Amp coulpled with methyl on Tmp
0.46444	939.35025	(944) P-O-H bend on Amp
		(936) P-O-H bend on Amp coupled with P-O stretch on Amp
		(847) symmetric double P-O on Amp
0.30809	885.362	(806) P-O-H bend on Amp (H- bonded to phosphate on Tmp)
0.36898	771.39241	(786) C-O-H bend on sugar of Tmp (H- bonded to own phosphate)
0.10487	649.79735	(648) N-H out of plane bend on Tmp

dGMP-dCMP

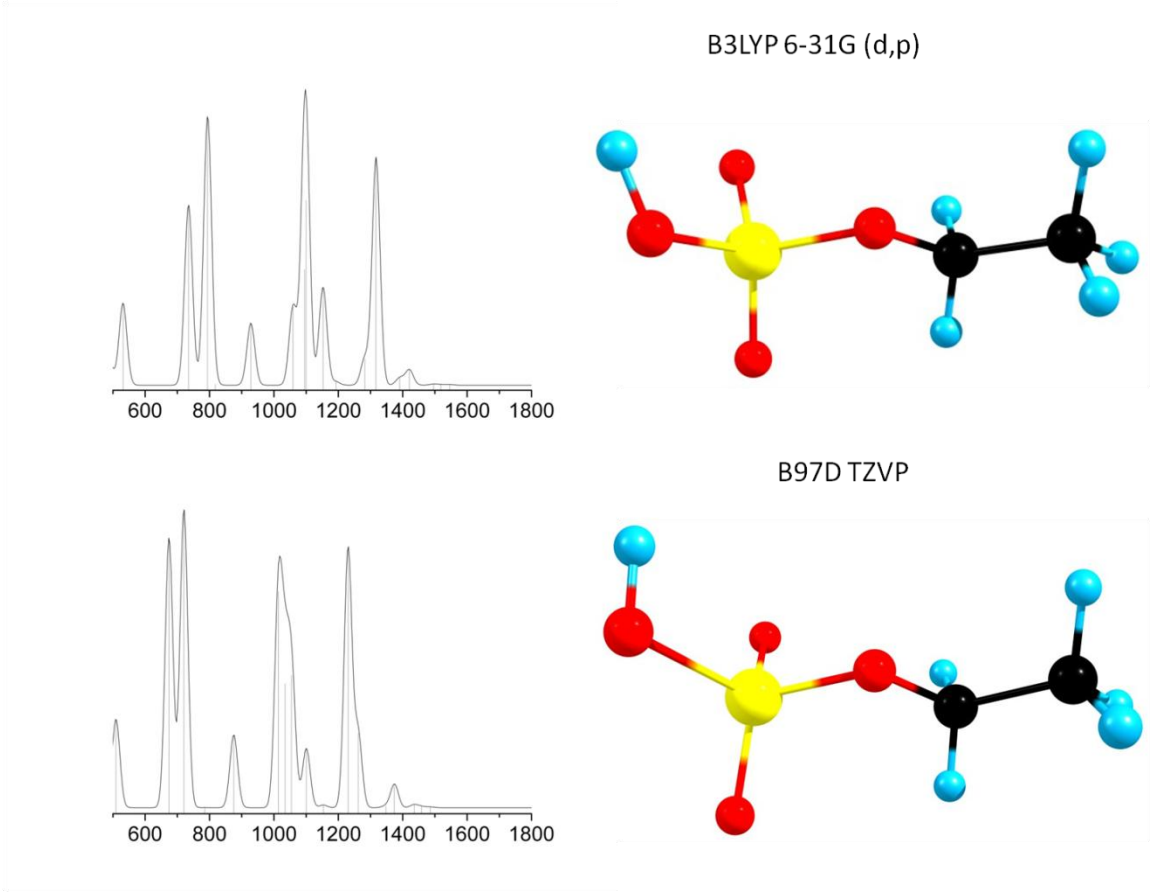
Experiment		Calculation
Intensity	Frequency (cm^{-1})	
0.27299	1702.69124	(1705) C6=O carbonyl stretch on dGmp
0.42662	1644.36833	(1660) C2=O carbonyl stretch on Cmp strongly coupled with N6-H inplane wag on dCmp and NH2 scissor on dGmp
		(1623) C2=O carbonyl stretch strongly coupled with C5-C6 stretch and inplane NH2 scissor on dCmp
0.29744	1589.92706	(1599) NH2 scissor strongly coupled to N-H wag all on dGmp
0.1714	1560.41419	
0.20305	1531.9826	
0.16401	1478.134	(1485) C8-N7 stretch on dGmp
		(1483) C4-N3 stretch on dCmp
0.14465	1415.94945	
0.29921	1326.70525	(1304) C8-H wag on dGmp
		(1279) C6-H in plane wag coulped to inplane NH2 movement on dCmp
0.43274	1257.41199	(1204) double P-O-H bend on dGmp
0.32829	1234.39724	(1129) asymmetric O-P-O with free P-O-H bend on dCmp
0.4	1129	(1059) free P-O-H bend on dCmp
0.64644	1090.79623	(1050) free PO-H bend on dCmp coupled with P=O stretch and phosphate bound C-O all on Cmp
0.6042	1037.23751	(1012) large symmetric P-O coupled to sugar ring mode all on dCmp
0.32325	977.30883	(984) C-O stretch connecting sugar to phosphate on dCmp
0.29329	937.61675	
0.27653	907.48199	
0.46411	876.21175	(794) P-O stretch on dCmp free P-O-H
0.58387	849.87201	(829) NH2 out of plane wag on dCmp (H bonded to C=O on dGmp)
		(794) P-O stretch on dCmp free P-O-H
0.28266	785.04782	(790) N-H out of plane bend on dCmp coupled to NH2 out of plane wag on dGmp

2. Experimental spectra of monomer and cluster plotted together (Top) [dAMP-dTMP]$^-$ (Bottom) [dGMP-dCMP]$^-$



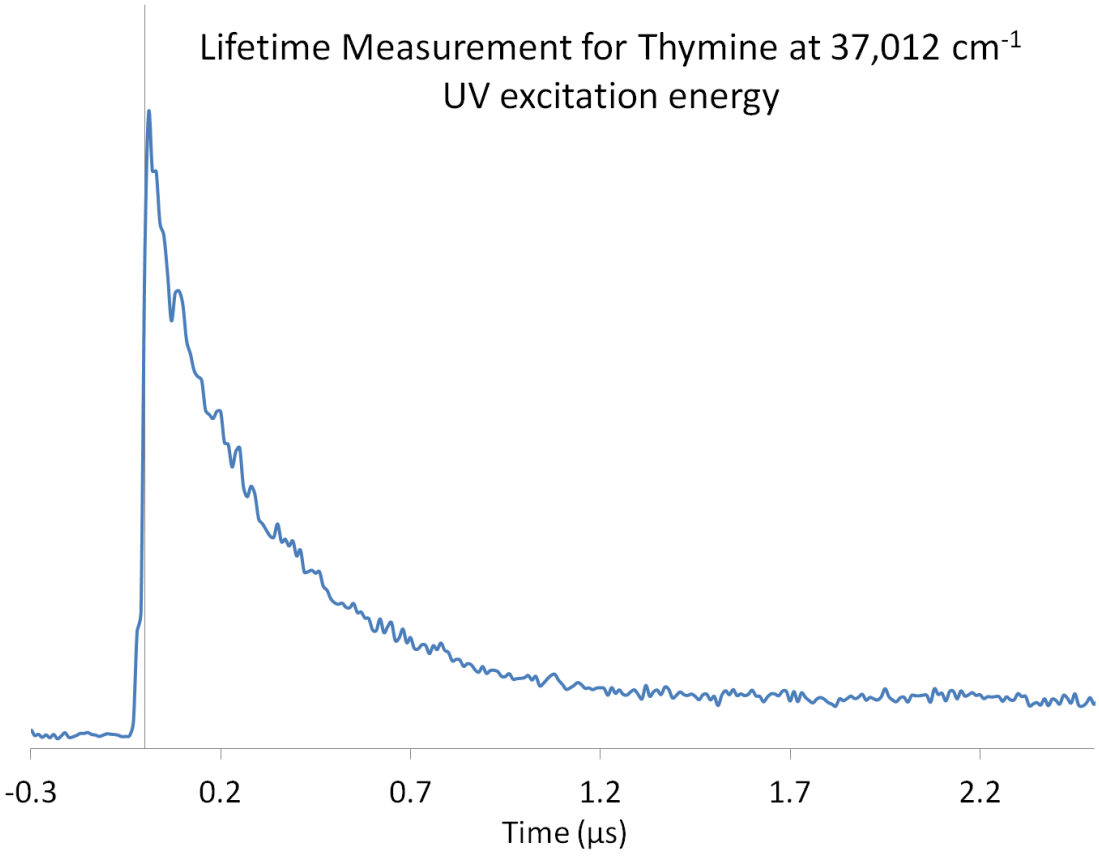

10
9
8
7
6
5
4
3
2
1
500
700
900
1100
1300
1500
1700
Wavenumber (cm-1)

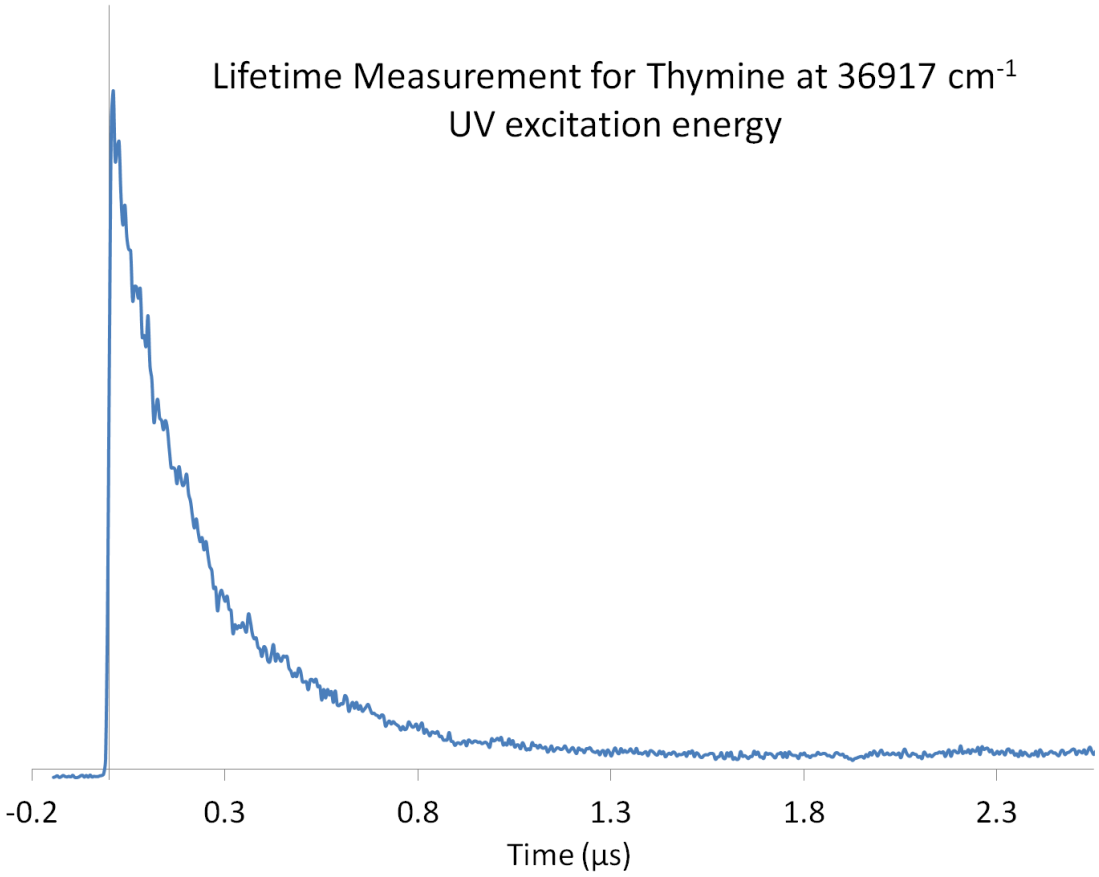
3. Calculated spectrum for ethylphosphate

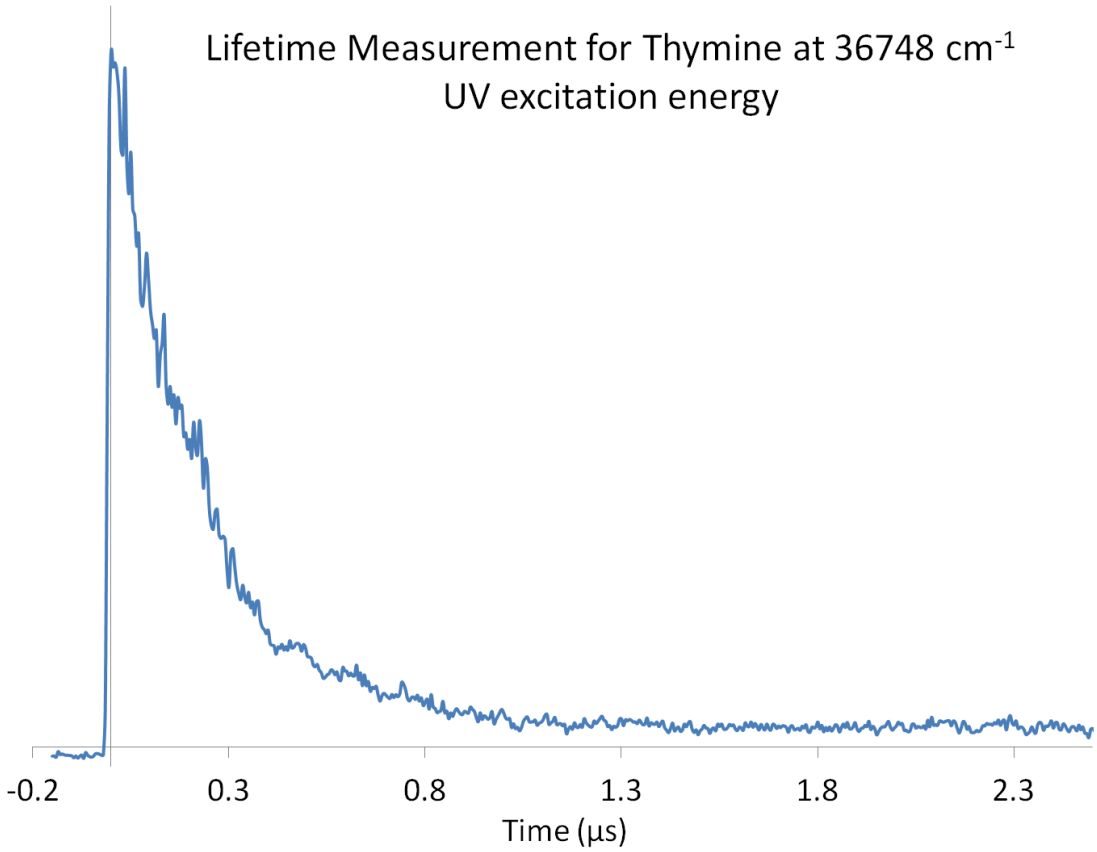


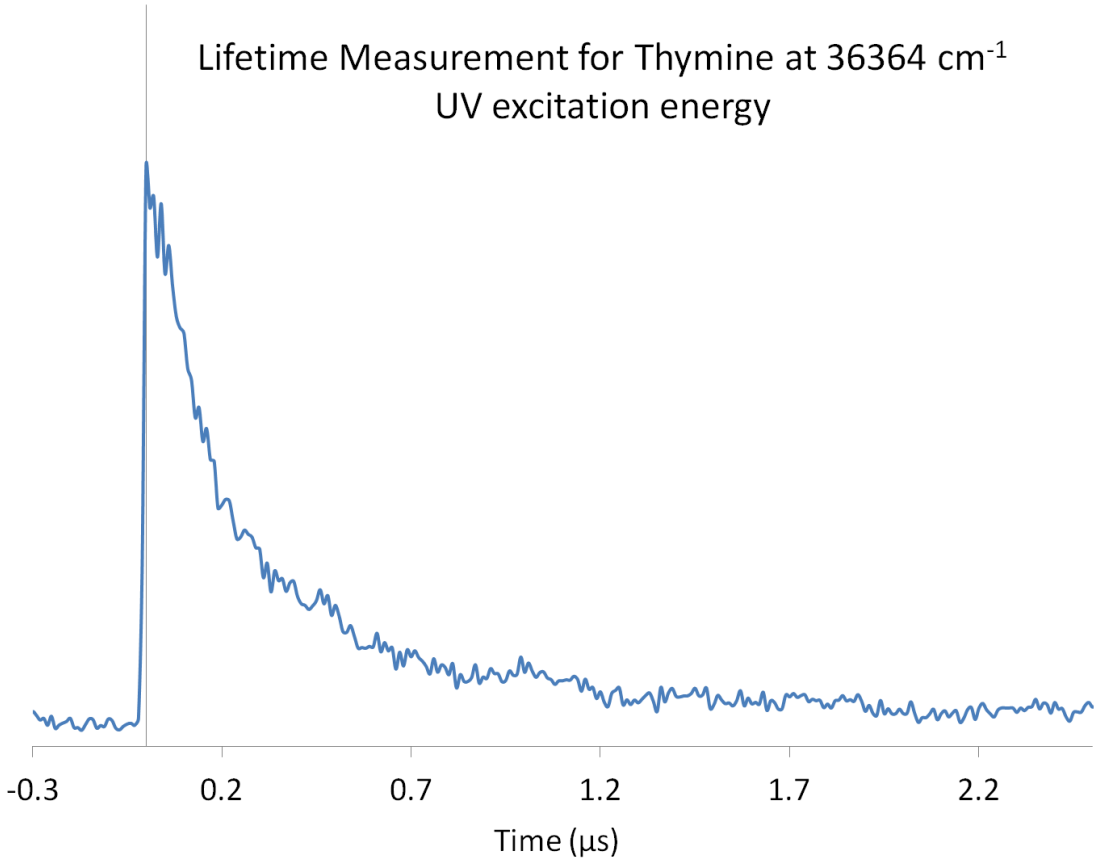
Appendix II

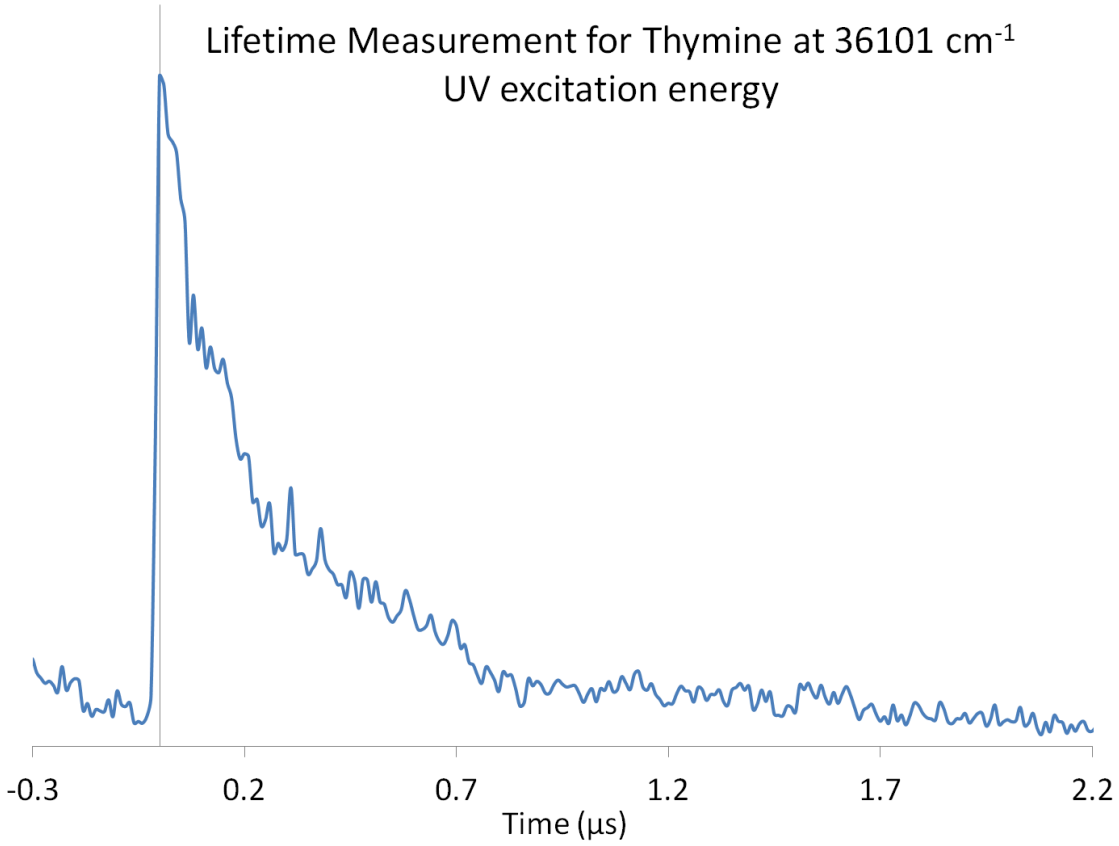
1. Lifetime Measurements for Thymine and Uracil

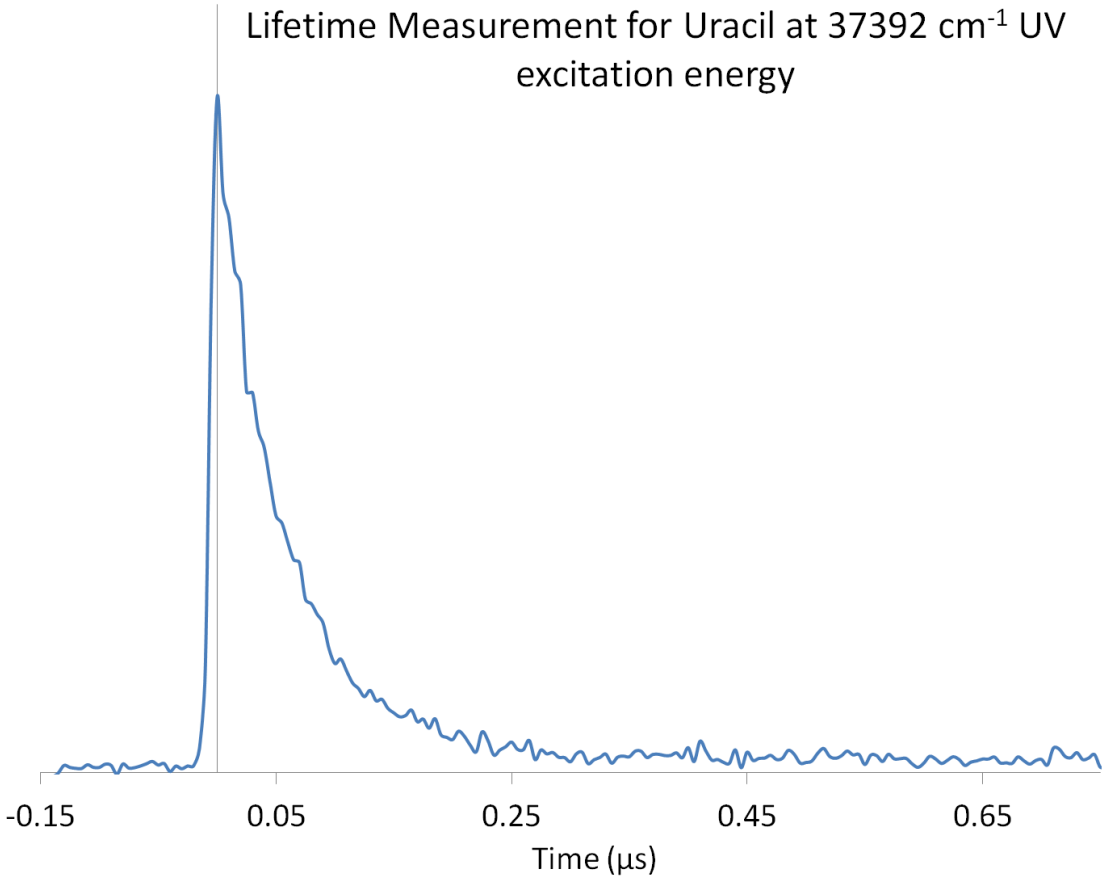


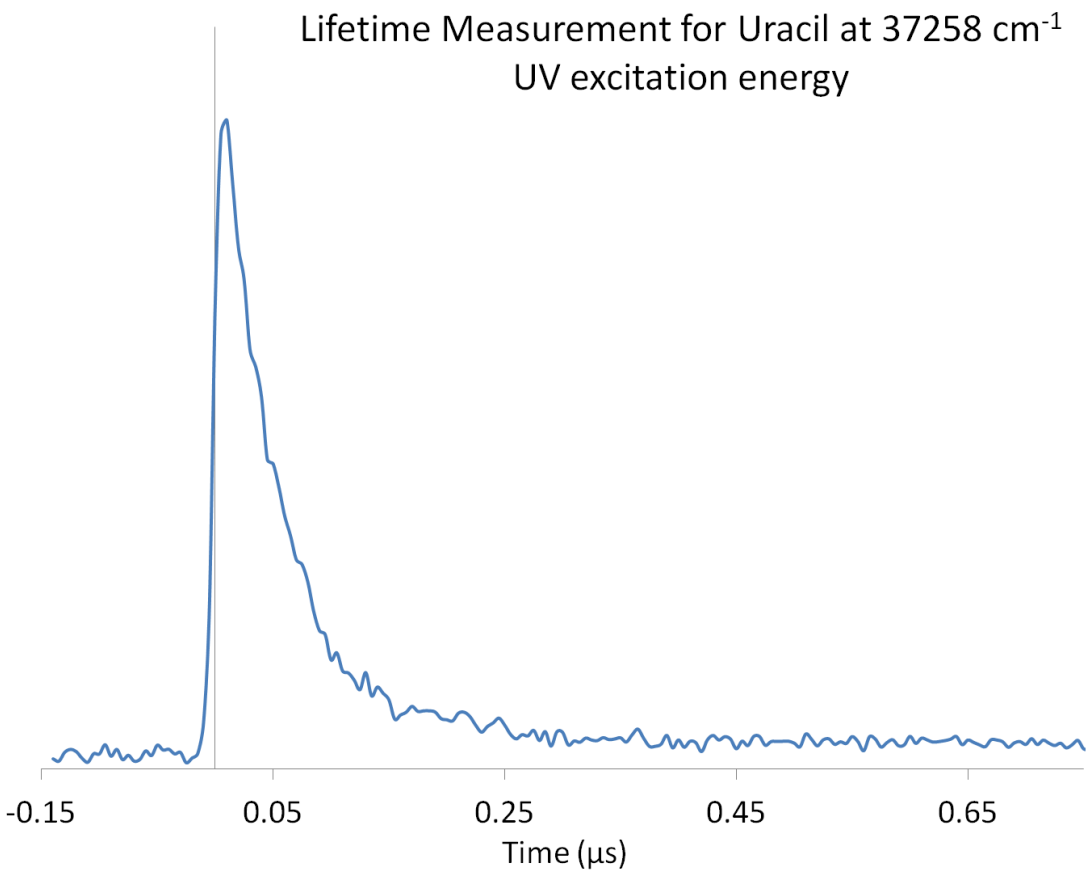

Lifetime Measurement for Thymine at 36917 cm-1
UV excitation energy
-0.2
0.3
0.8
1.3
1.8
2.3
Time (μs)

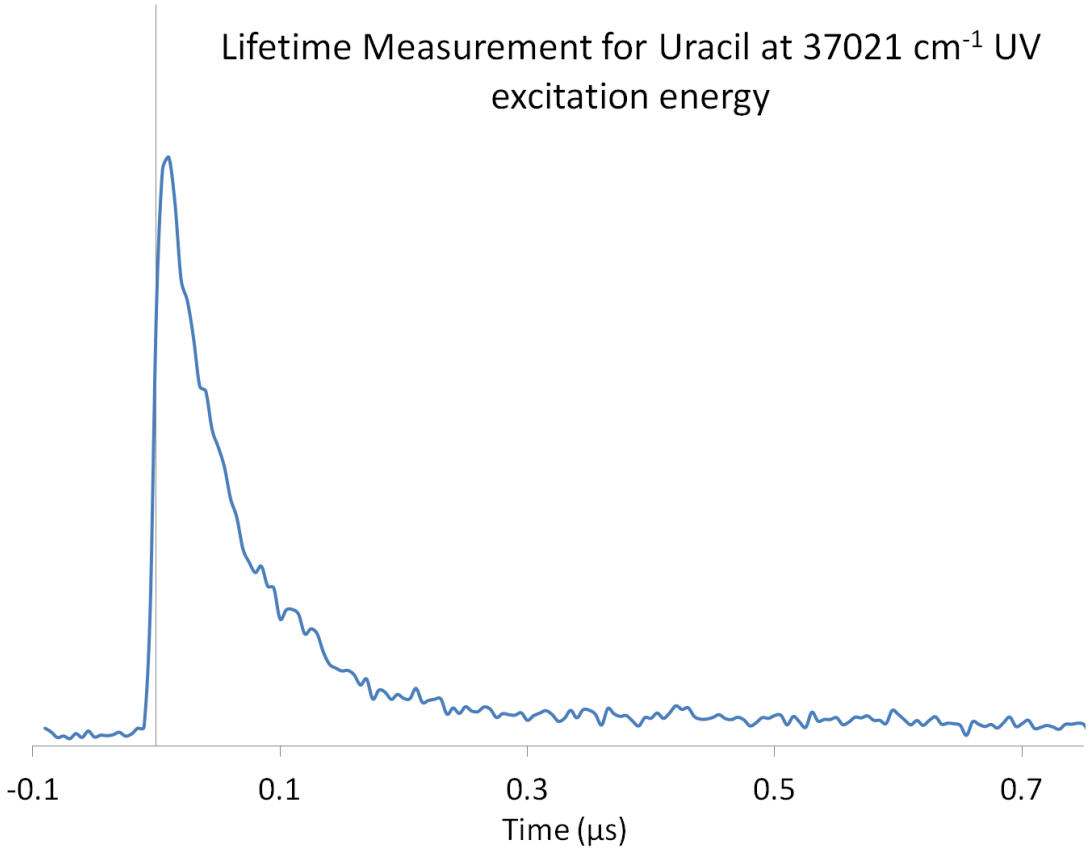

Lifetime Measurement for Thymine at 36748 cm-1
UV excitation energy
-0.2
0.3
0.8
1.3
1.8
2.3
Time (µs)

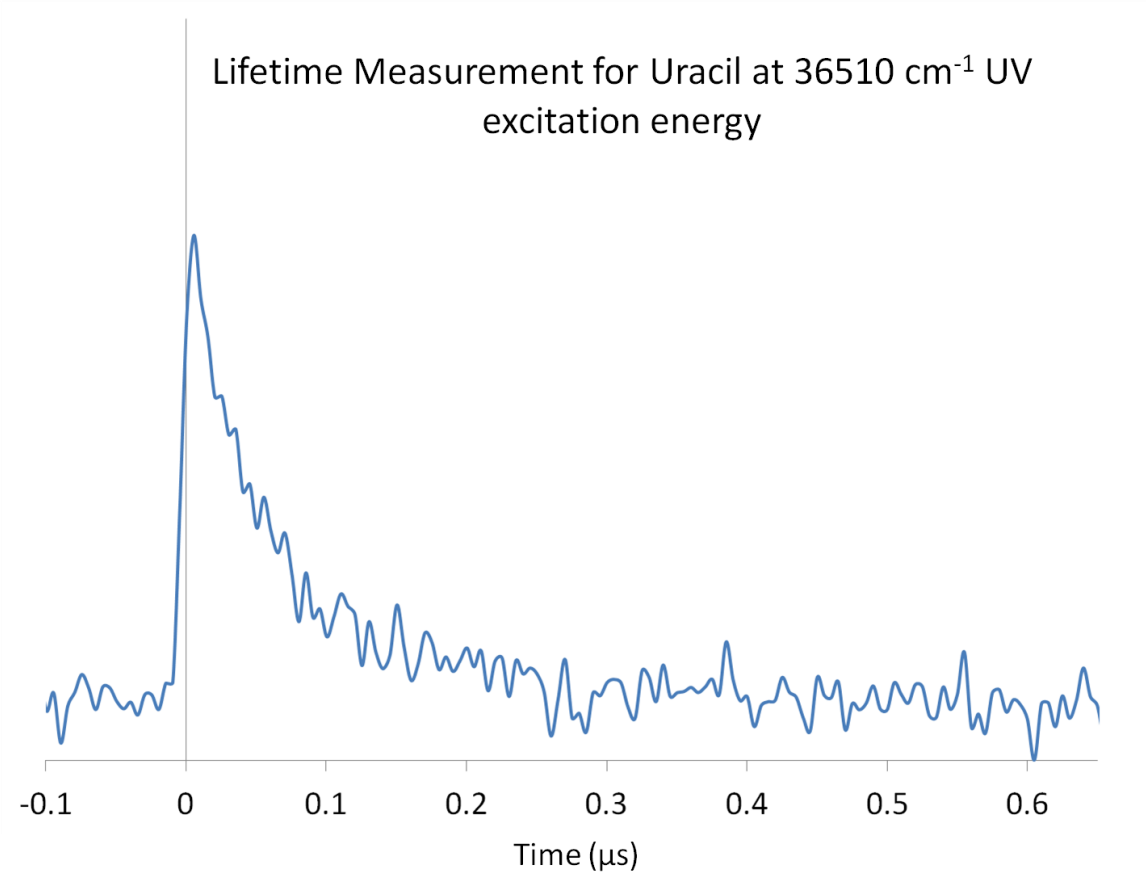

Lifetime Measurement for Thymine at 36364 cm-1
UV excitation energy
-0.3
0.2
0.7
1.2
1.7
2.2
Time (µs)

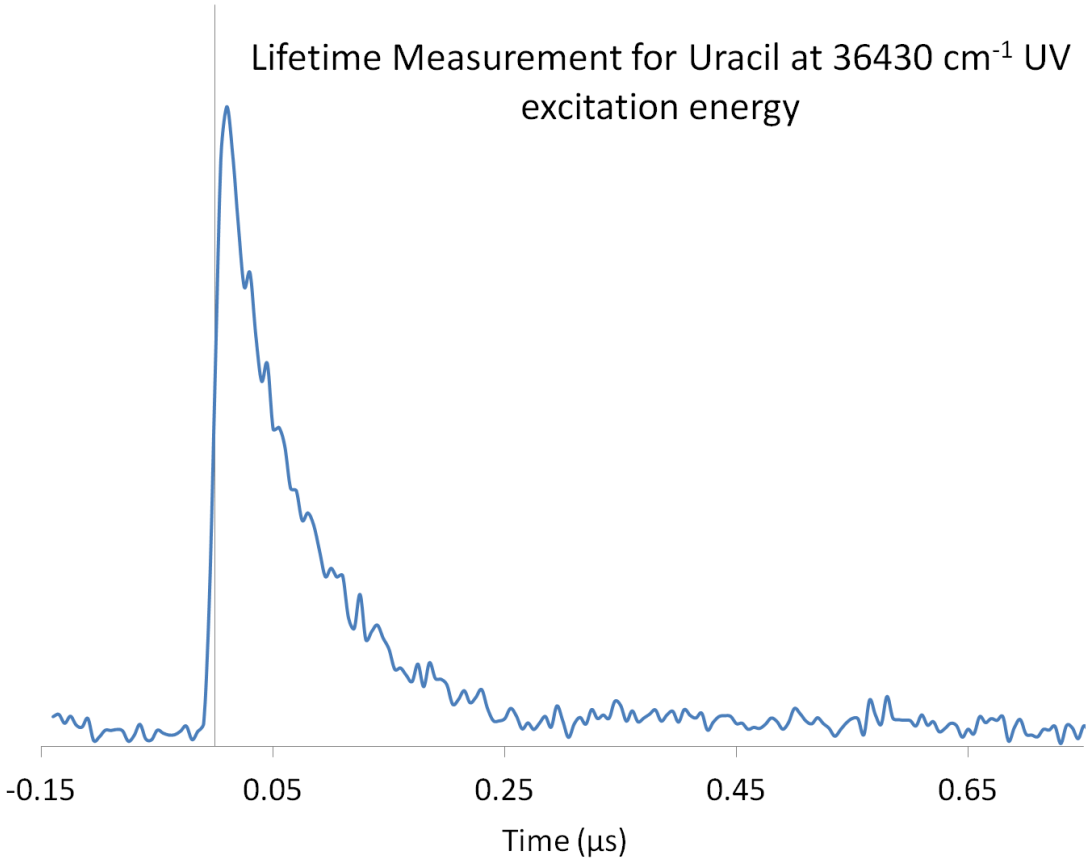

Lifetime Measurement for Thymine at 36101 cm-1
UV excitation energy
-0.3
0.2
0.7
1.2
1.7
2.2
Time (μs)


Lifetime Measurement for Uracil at 37392 cm-1 UV excitation energy
-0.15
0.05
0.25
0.45
0.65
Time (µs)


Lifetime Measurement for Uracil at 37258 cm-1
UV excitation energy
-0.15
0.05
0.25
0.45
0.65
Time (µs)


Lifetime Measurement for Uracil at 37021 cm-1 UV excitation energy
-0.1
0.1
0.3
0.5
0.7
Time (µs)


Lifetime Measurement for Uracil at 36510 cm-1 UV excitation energy
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
Time (μs)


Lifetime Measurement for Uracil at 36430 cm-1 UV excitation energy
-0.15
0.05
0.25
0.45
0.65
Time (µs)

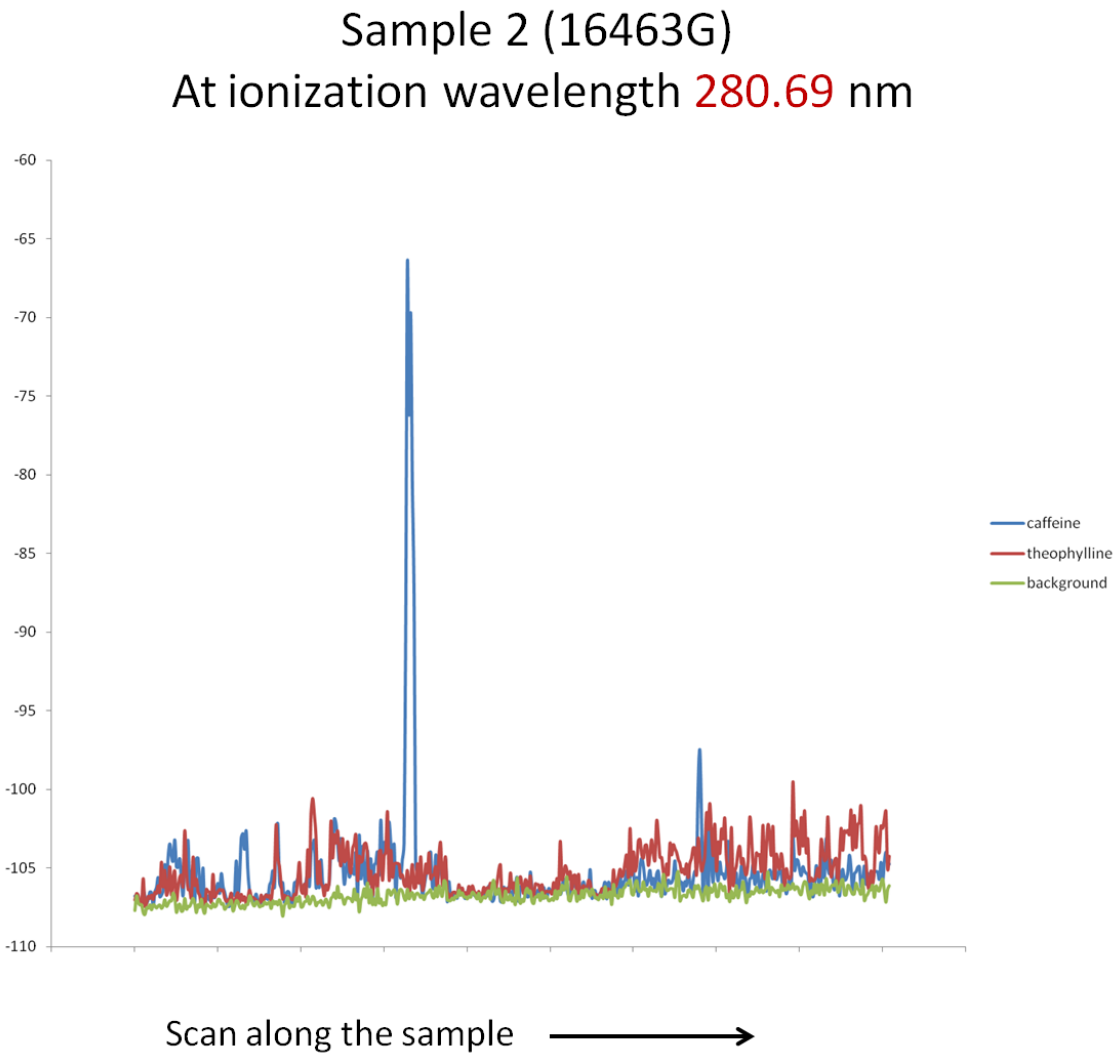
Appendix III

1. Tables for identification of xanthine stimulants in pottery sherds

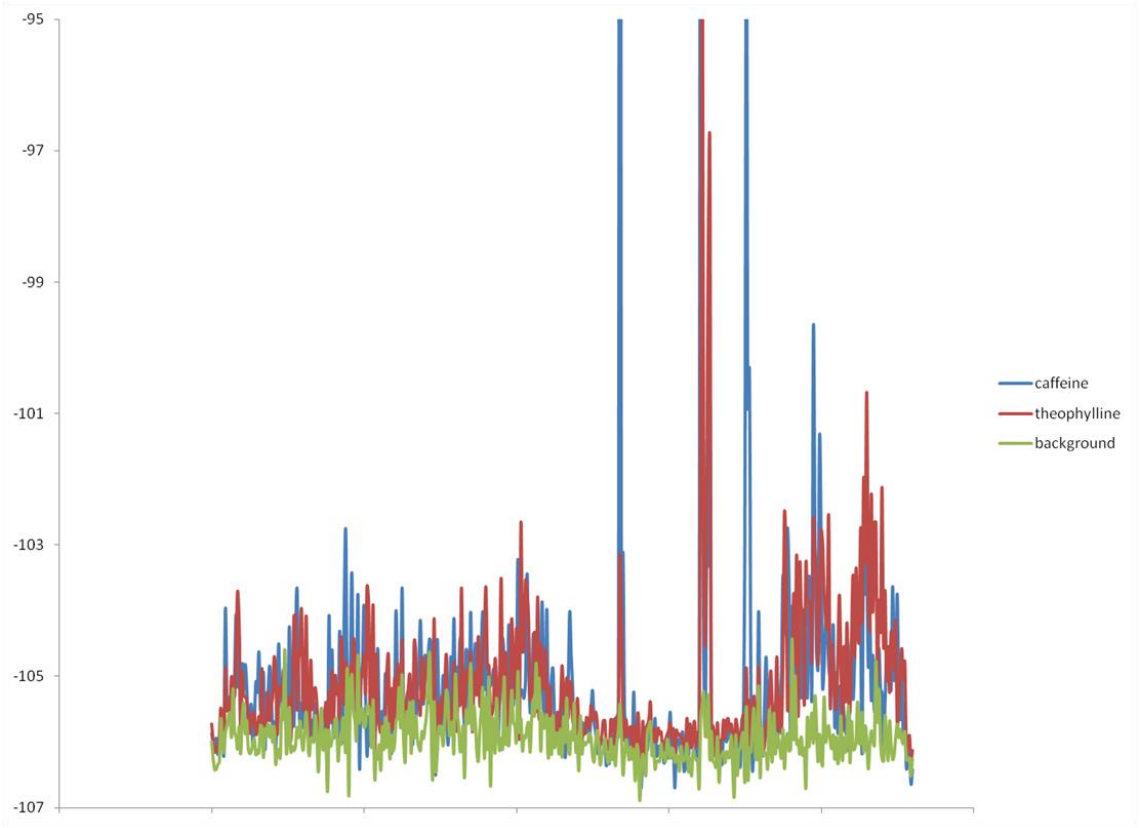
Sample #	Sherd ID #	Mass used for extract (g)	Where	Identification of theobromine by extraction	Identification of theobromine by direct desorption	identification of theophylline by direct desorption	Identification of caffeine by direct desorption
1	106 578 1-034	0.4003	whole strata	positive	positive	positive	positive
2	16463G	0.4001	whole strata	positive	-	positive	positive
3	106 2009 8-046	0.4003	whole strata	positive	-	positive	positive
4	105 4209 8-002	0.4000	whole strata	positive	positive	positive	positive
5	16535Z	0.4003	whole strata	positive	-	positive	positive
6	16359AB	0.4001	mostly inner	positive	-	positive	positive
7	13394AO	0.4006	whole strata	positive	positive	positive	positive
8	14321V	0.4003	lighter colored inside layer	positive	-	positive	positive
9	14328K	0.4000	whole strata	positive	-	positive	positive
10	143325U	0.4004	whole strata	positive	-	positive	positive
11	14320M	0.4005	whole strata	negative	-	negative	negative
12	14319O	0.4004	some whole w extra inside	positive	-	positive	positive
13	14342V	0.4001	whole strata	positive	-	positive	positive
14	8 outside	0.3585	outside	trace	-	-	-

Sample #	Sherd ID #	Identification of theobromine by direct desorption	identification of theophylline by direct desorption	Identification of caffeine by direct desorption
1	11 Fv 47 82-273	trace	negative	positive
2	11 Fv 47 F.52.21 ZB NE1/2	positive	positive	positive
3	CW Cooper Featrue 12 F. 12-8 Zone A SE1/2	positive	trace	positive
4	11 Fv 47 lot 31 82A feature 61	negative	trace	positive
5	Professor Greg Wilson	trace	positive	positive
6	11 Fv 47 F.39 Profile wall	trace	positive	positive
7	11 Fv 4n Lot 273 82A	positive	positive	positive

2. Scans of the Maya pottery sherds showing positive identification of xanthine stimulants by direct desorption

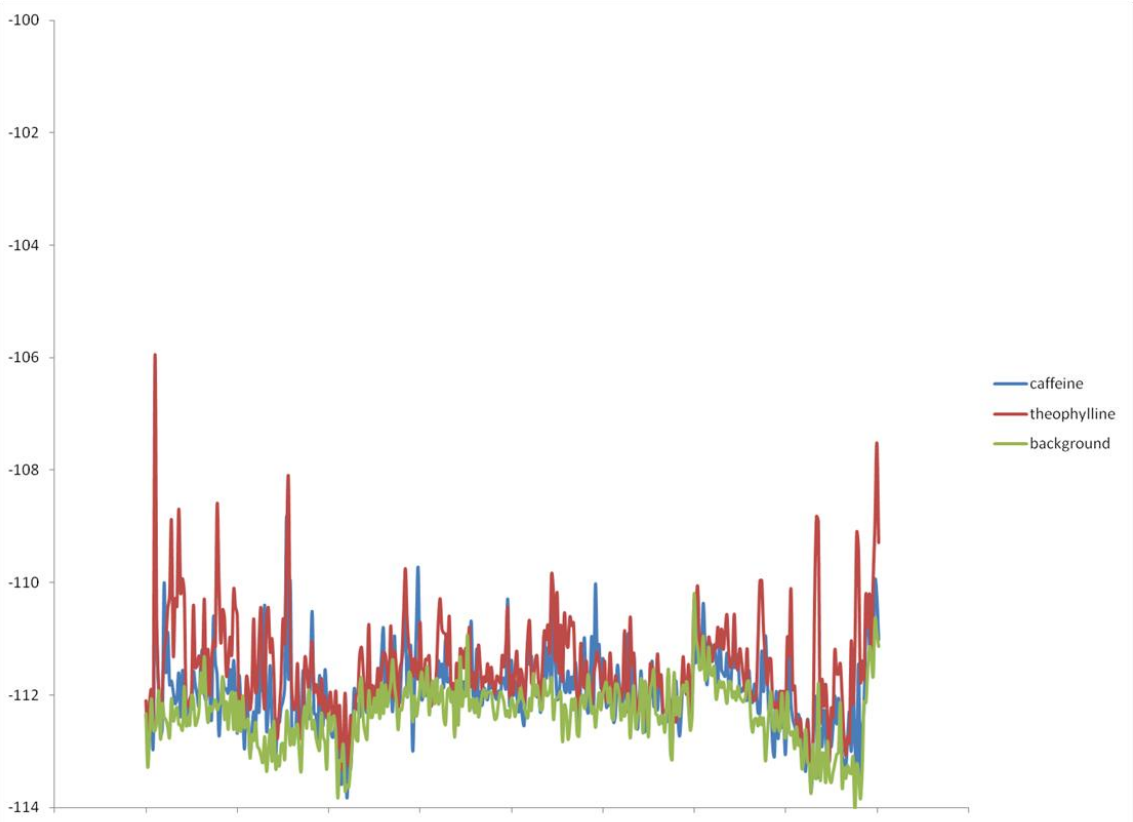


Sample 3 (106 2009 8-046)
At ionization wavelength 280.69 nm



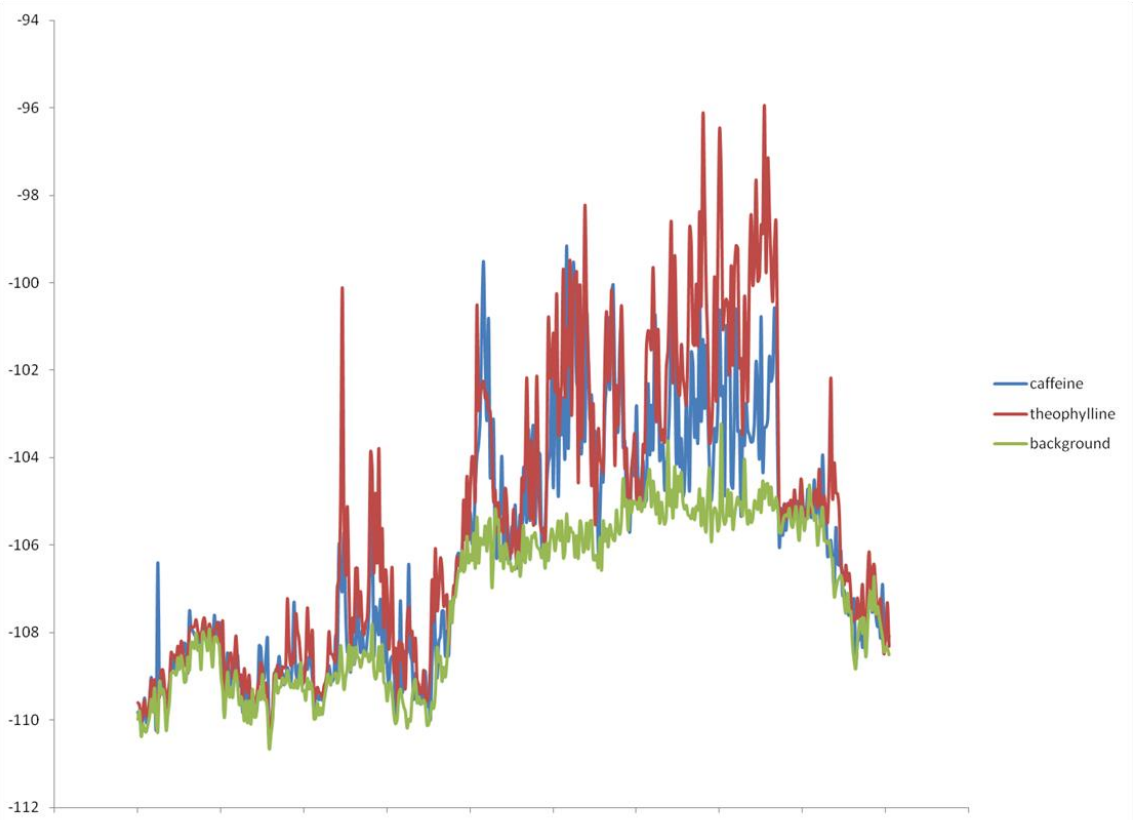
Scan along the sample →

Sample 4 (105 4209 8-002)
At ionization wavelength 280.69 nm



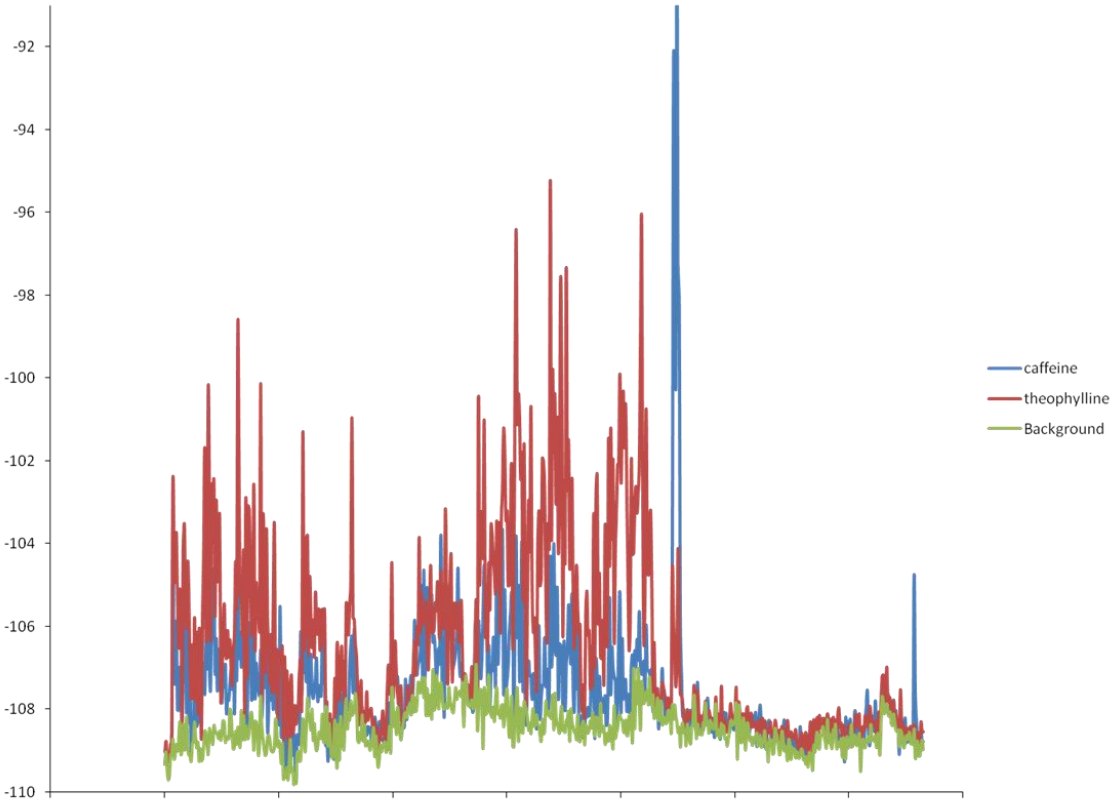
Scan along the sample →

Sample 5 (16535Z)
At ionization wavelength 280.69 nm



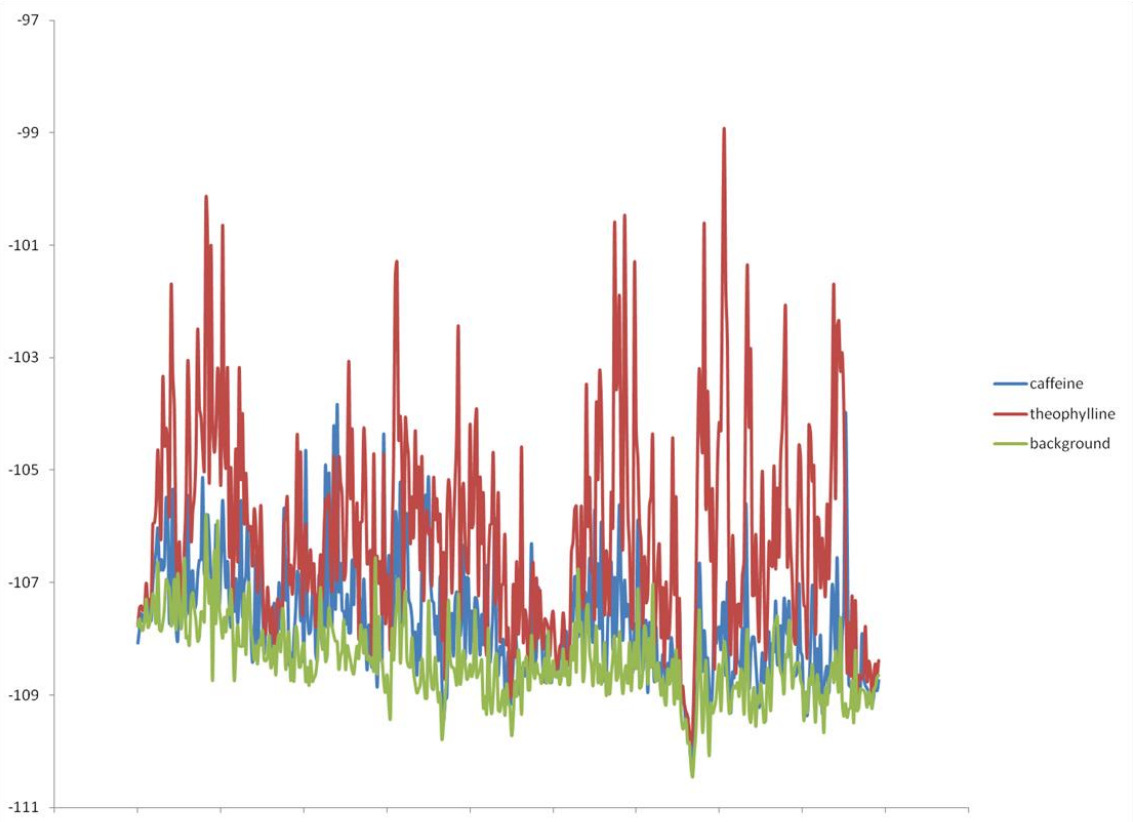
Scan along the sample →

Sample 6 (16359AB)
At ionization wavelength 280.69 nm



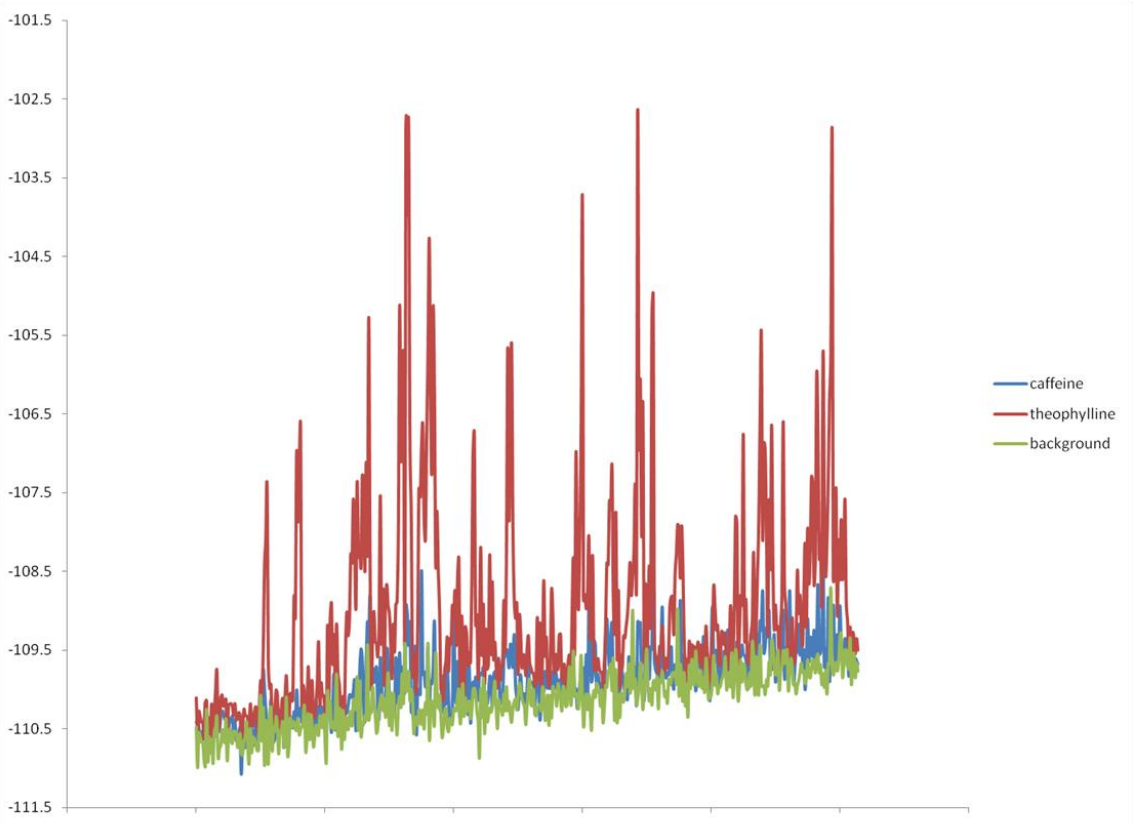
Scan along the sample →

Sample 8 (14321V)
At ionization wavelength 280.69 nm



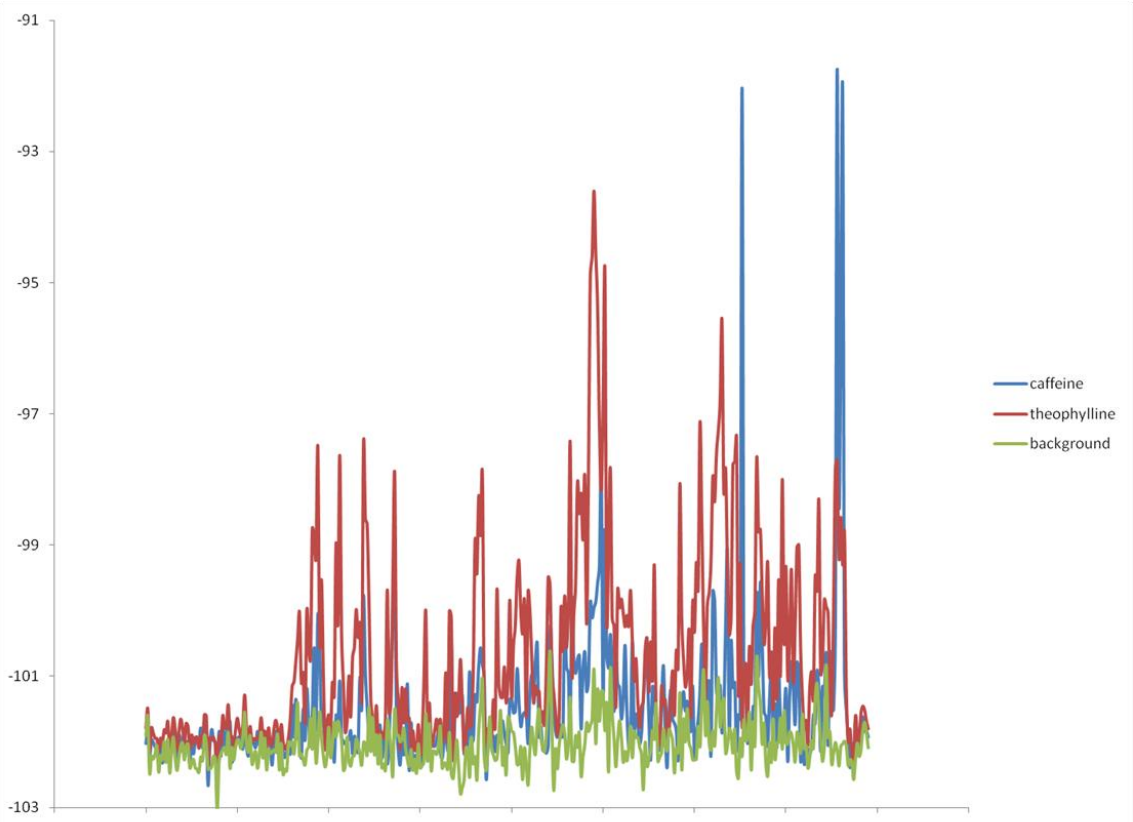
Scan along the sample →

Sample 9 (14328K)
At ionization wavelength 280.69 nm



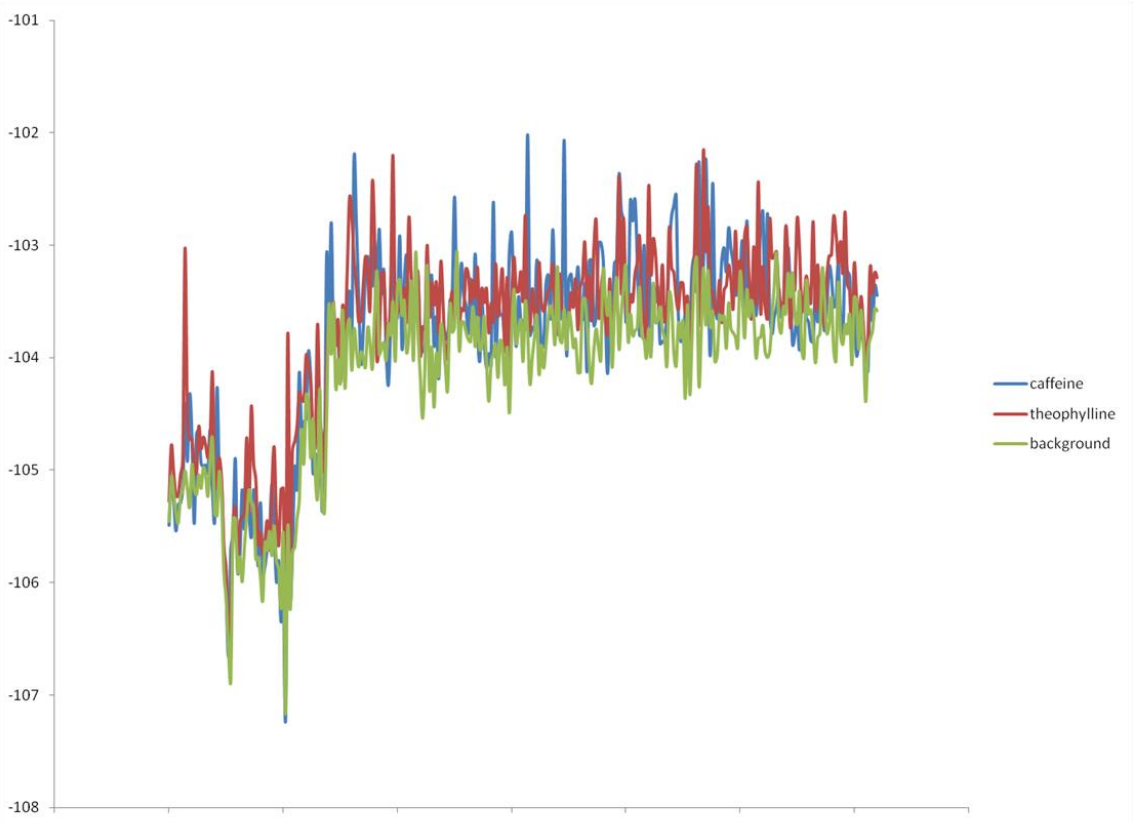
Scan along the sample →

Sample 10 (143325U)
At ionization wavelength 280.69 nm



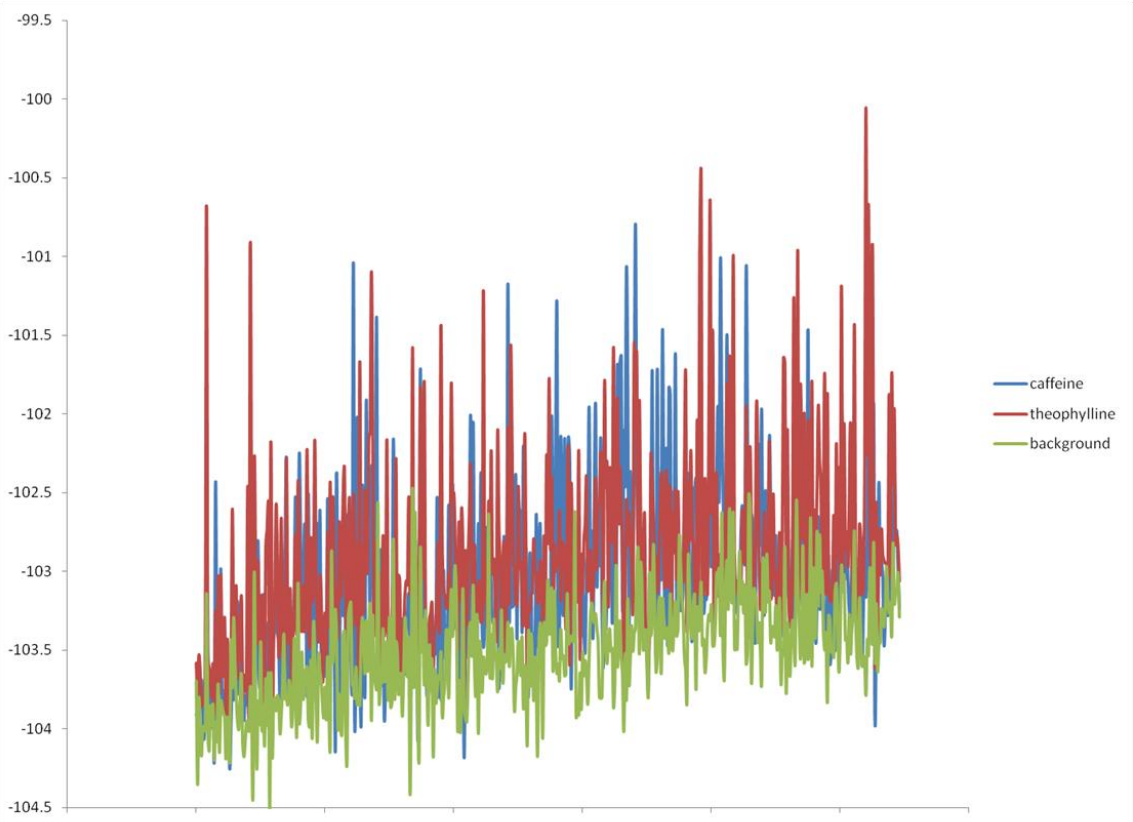
Scan along the sample →

Sample 11 (14320M)
At ionization wavelength 280.69 nm



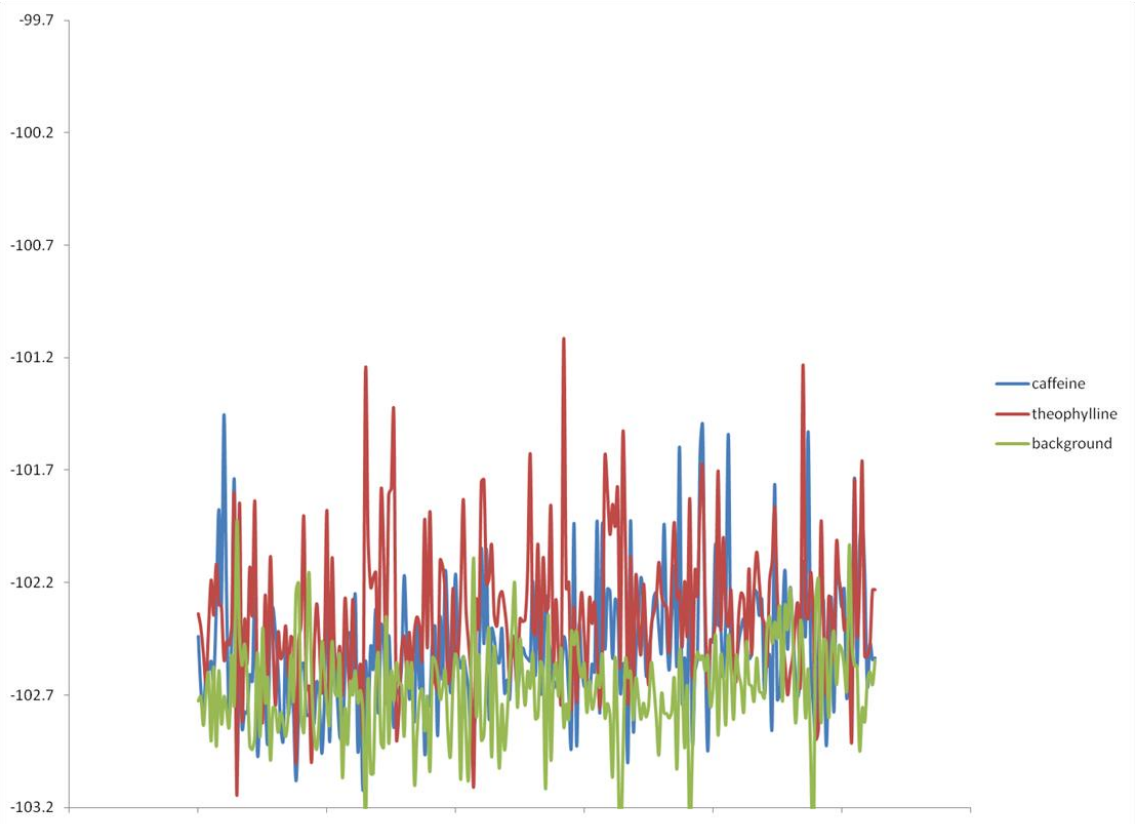
Scan along the sample →

Sample 12 (14319O)
At ionization wavelength 280.69 nm



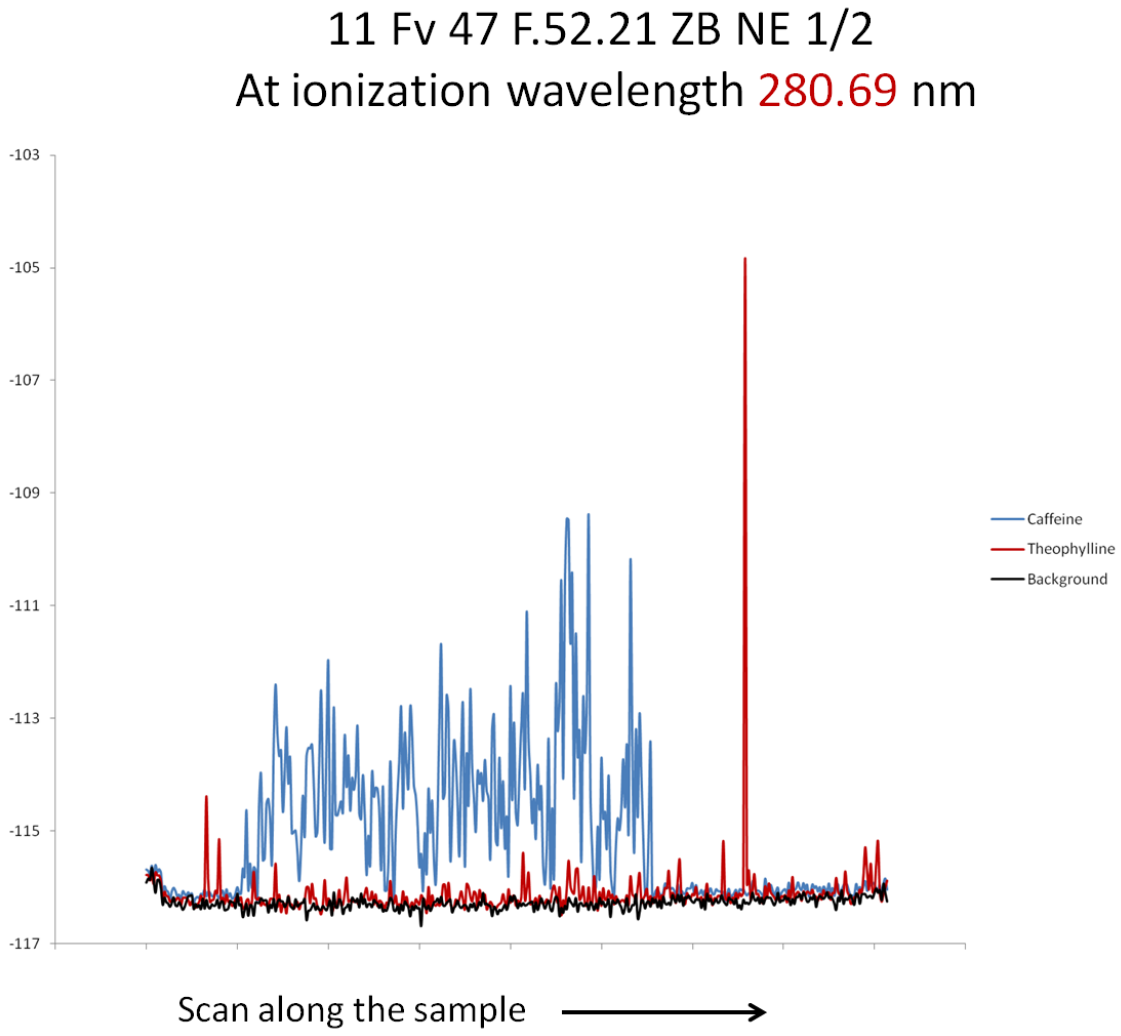
Scan along the sample →

Sample 13 (14342V)
At ionization wavelength 280.69 nm

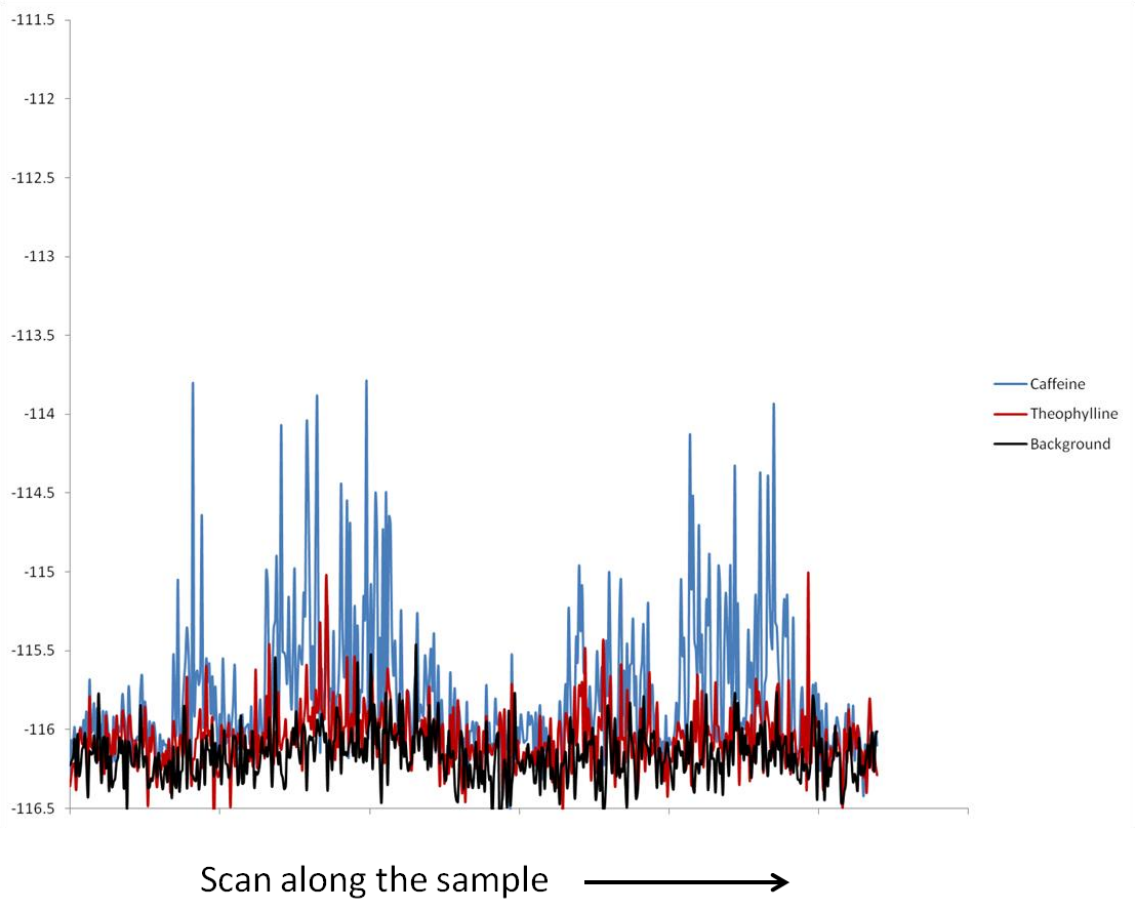


Scan along the sample →

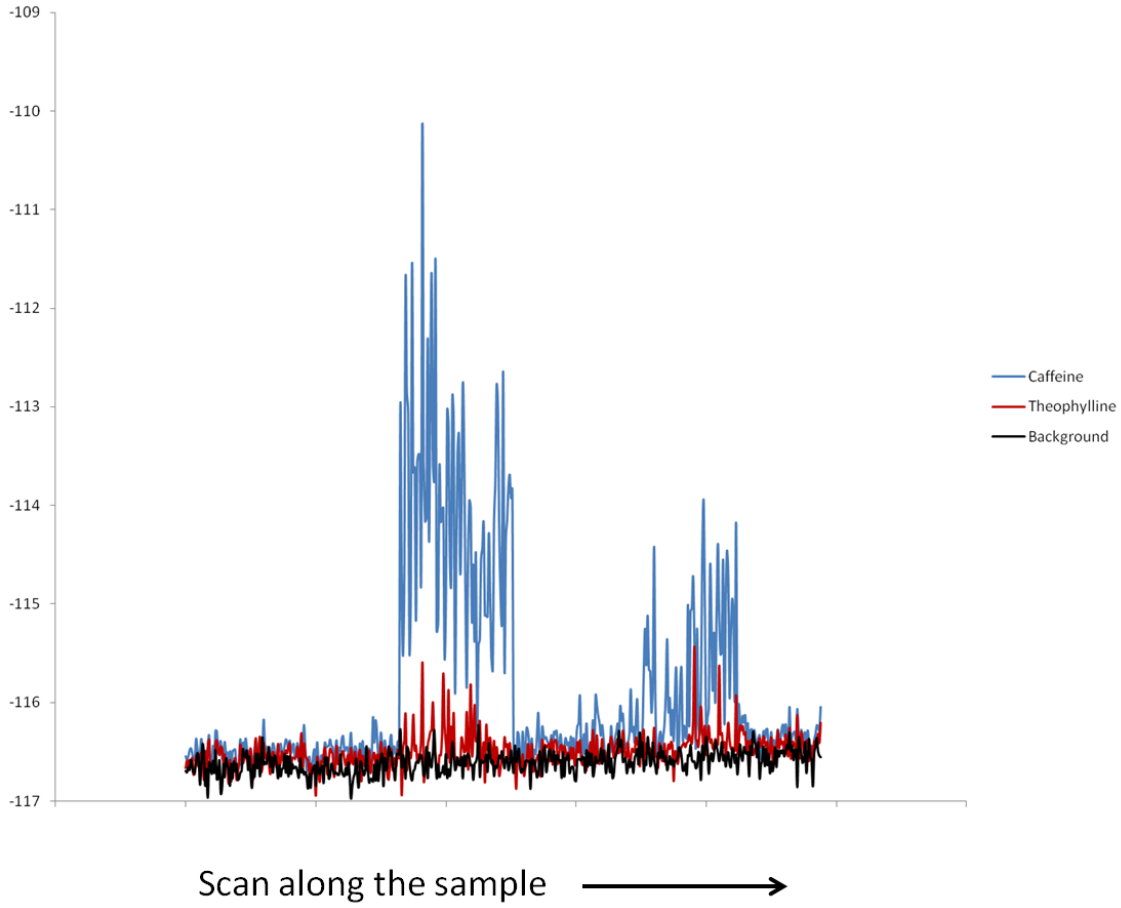
3. Scans of the Mississippi pottery sherds showing Positive identification of xanthine stimulants by direct desorption



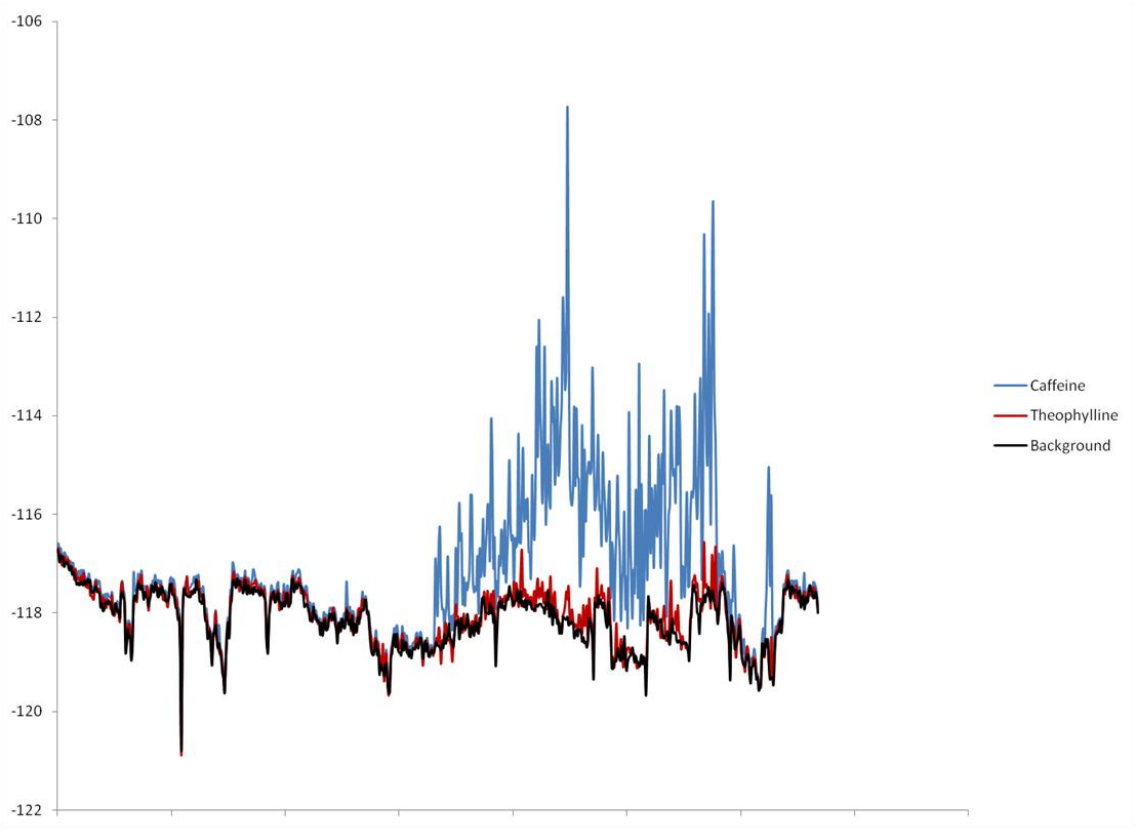
11Fv-47 lot 31 82A Feature 61
At ionization wavelength 280.69 nm



CW Cooper Feature 12 F.12-8 Zone A SE 1/2
At ionization wavelength 280.69 nm

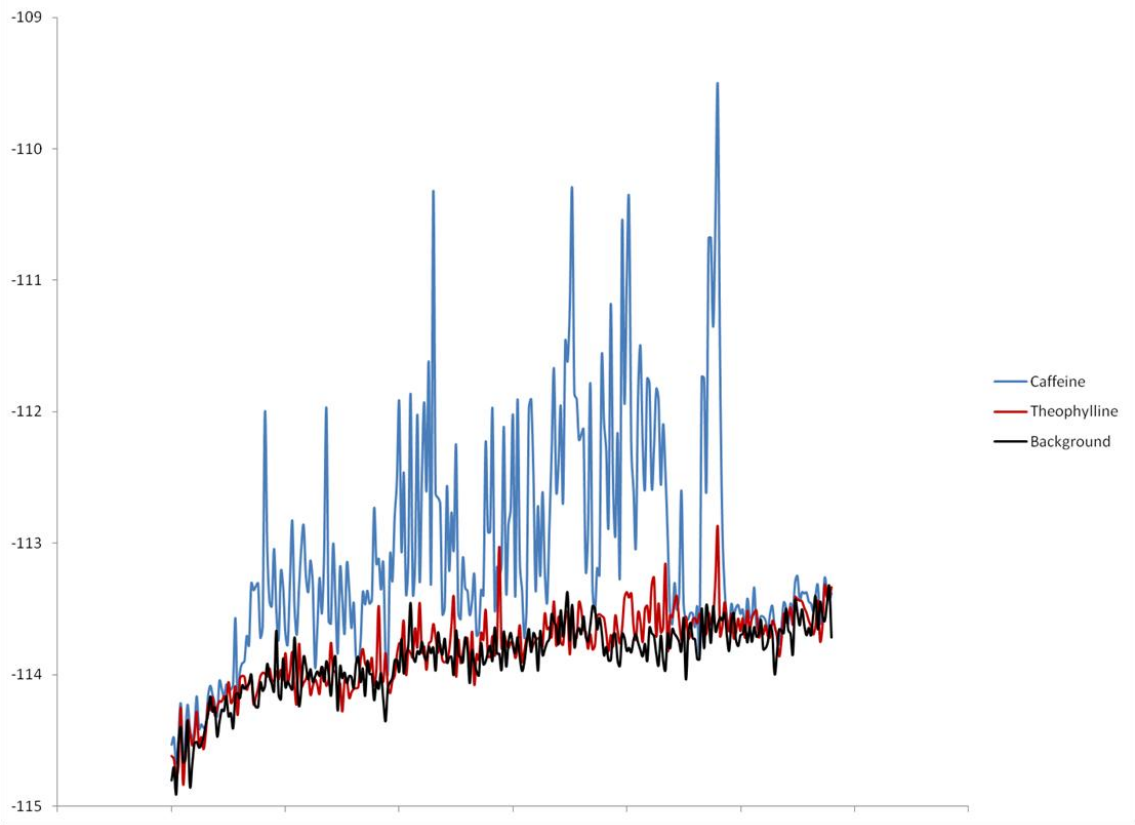


Professor Greg Wilson
At ionization wavelength 280.69 nm



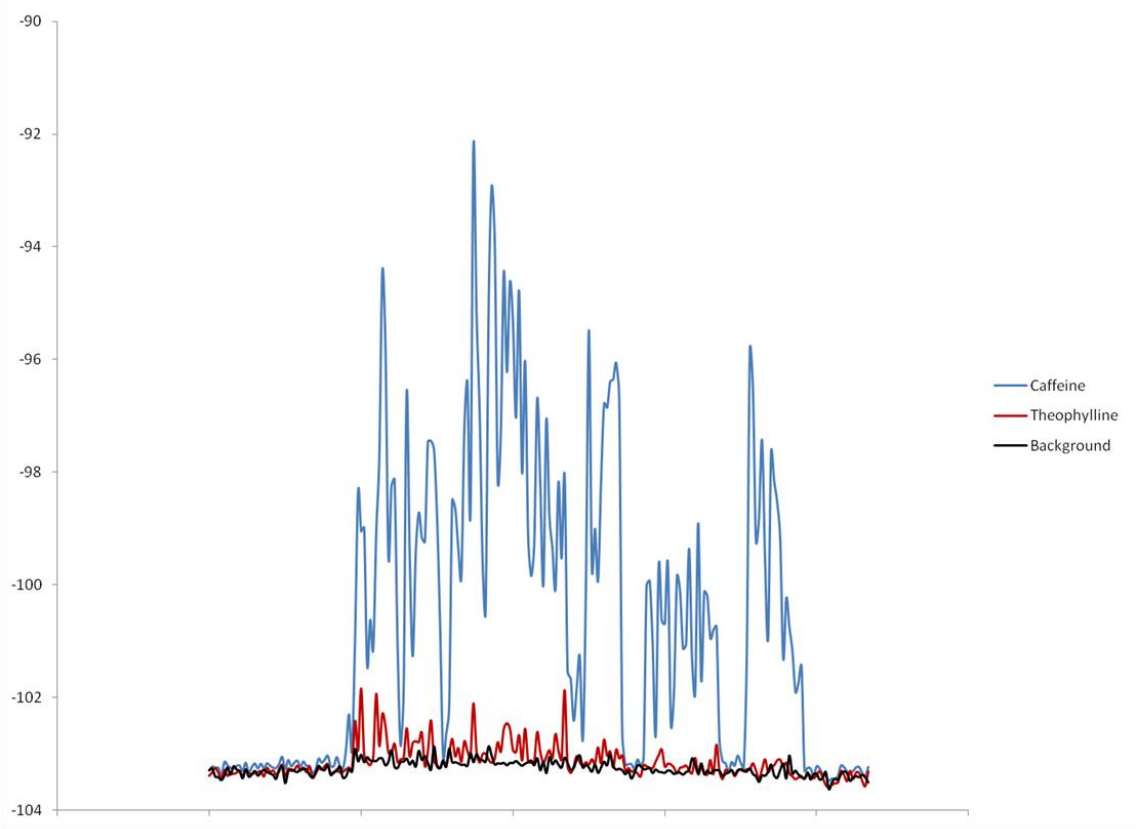
Scan along the sample →

11 Fv 47 F.39 Profile Wall
At ionization wavelength 280.69 nm



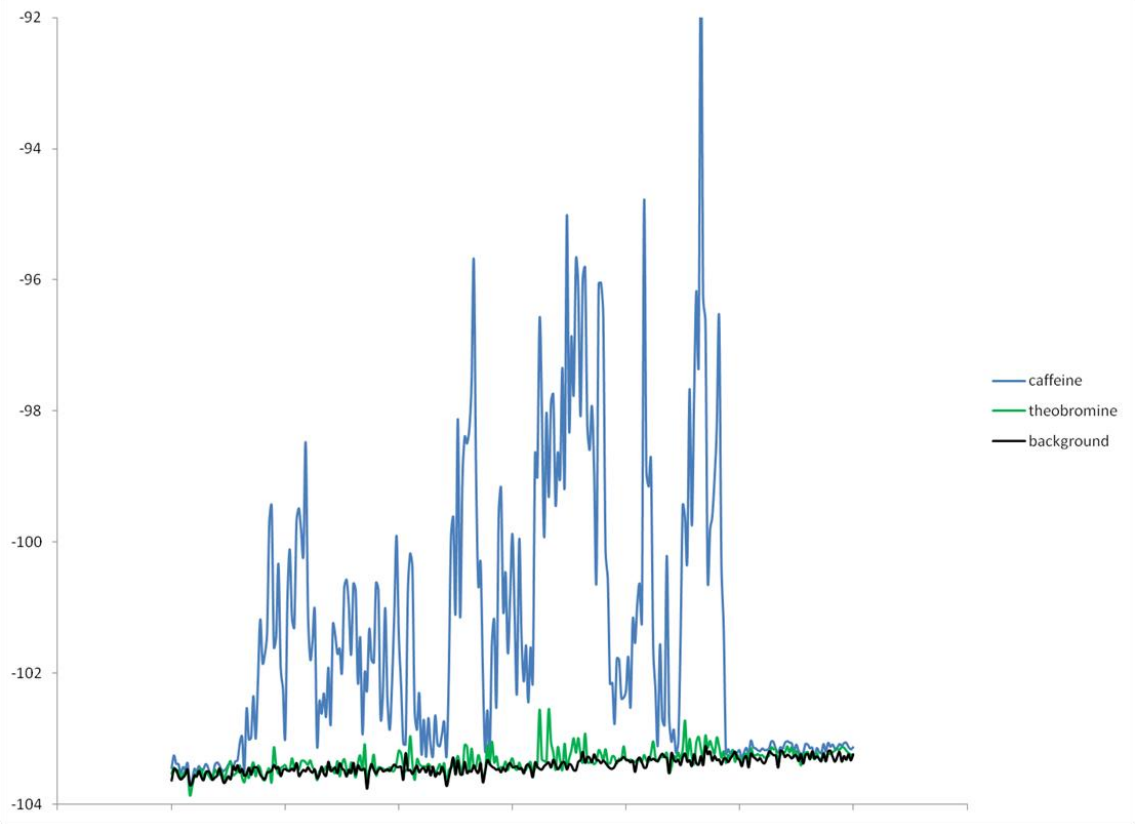
Scan along the sample →

11 Fv 4n lot 273 82A
At ionization wavelength 280.69 nm



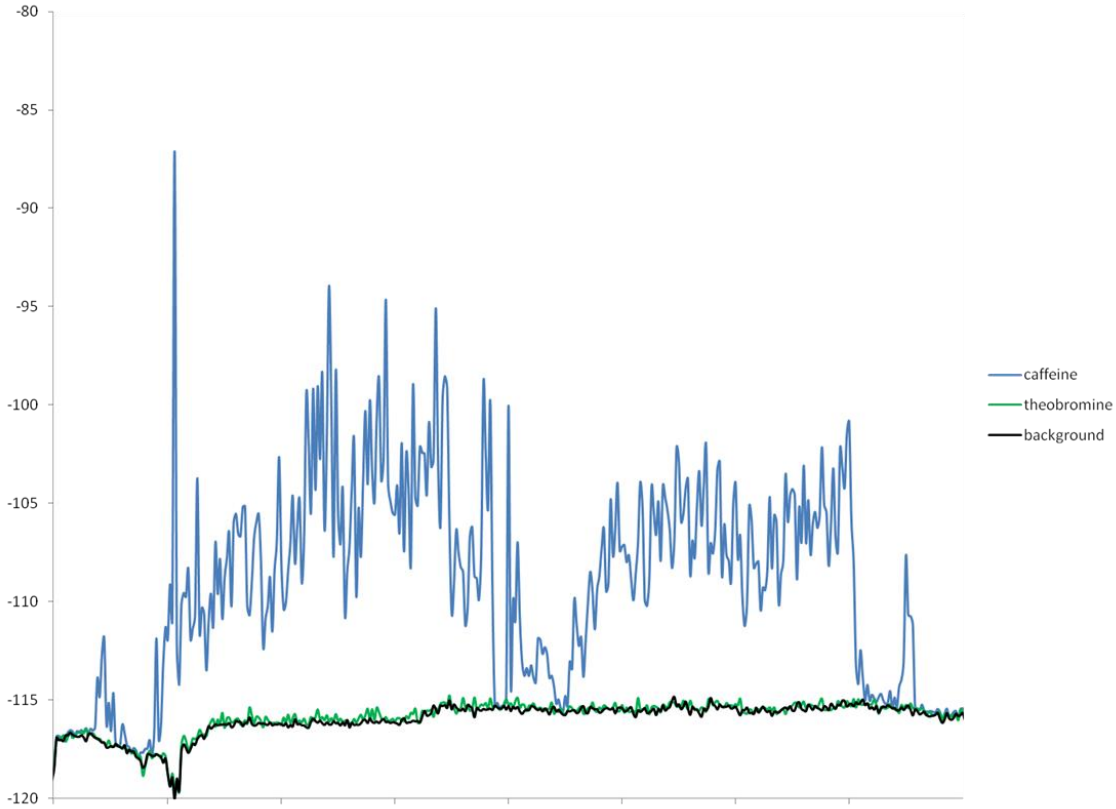
Scan along the sample →

11 Fv 4n lot 273 82A
At ionization wavelength 281.55 nm



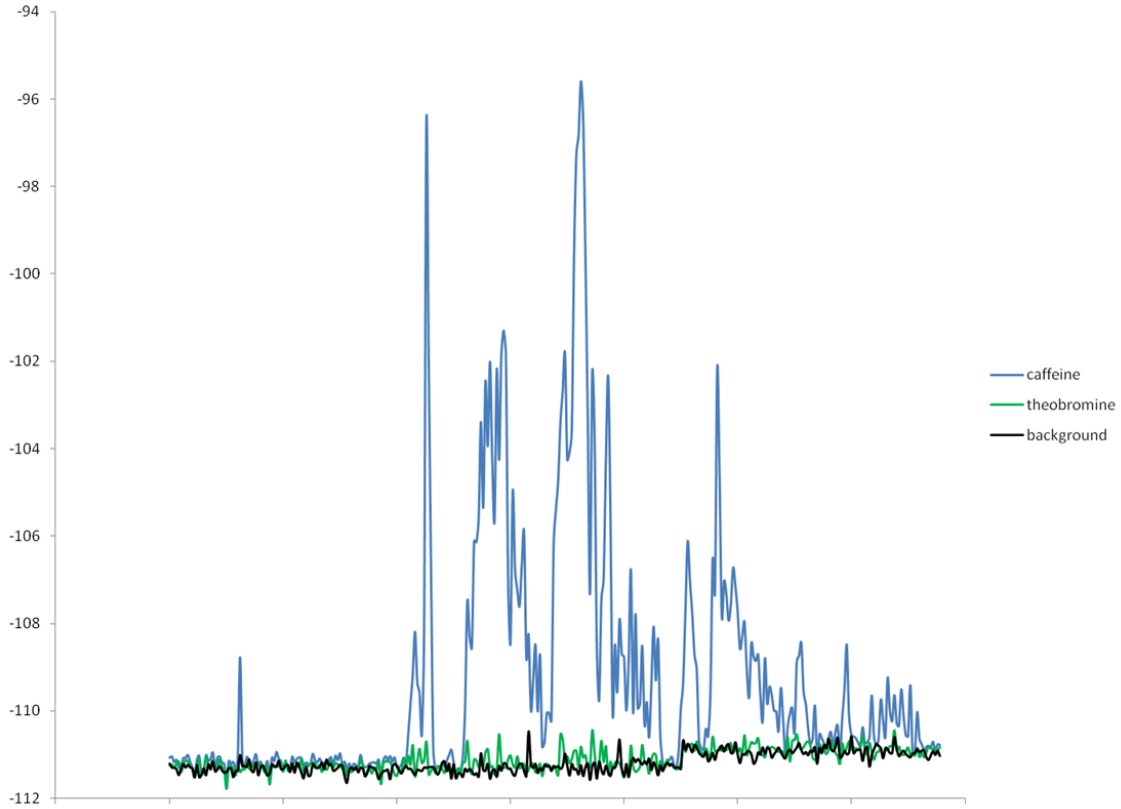
Scan along the sample →

Professor Greg Wilson
At ionization wavelength 281.55 nm



Scan along the sample →

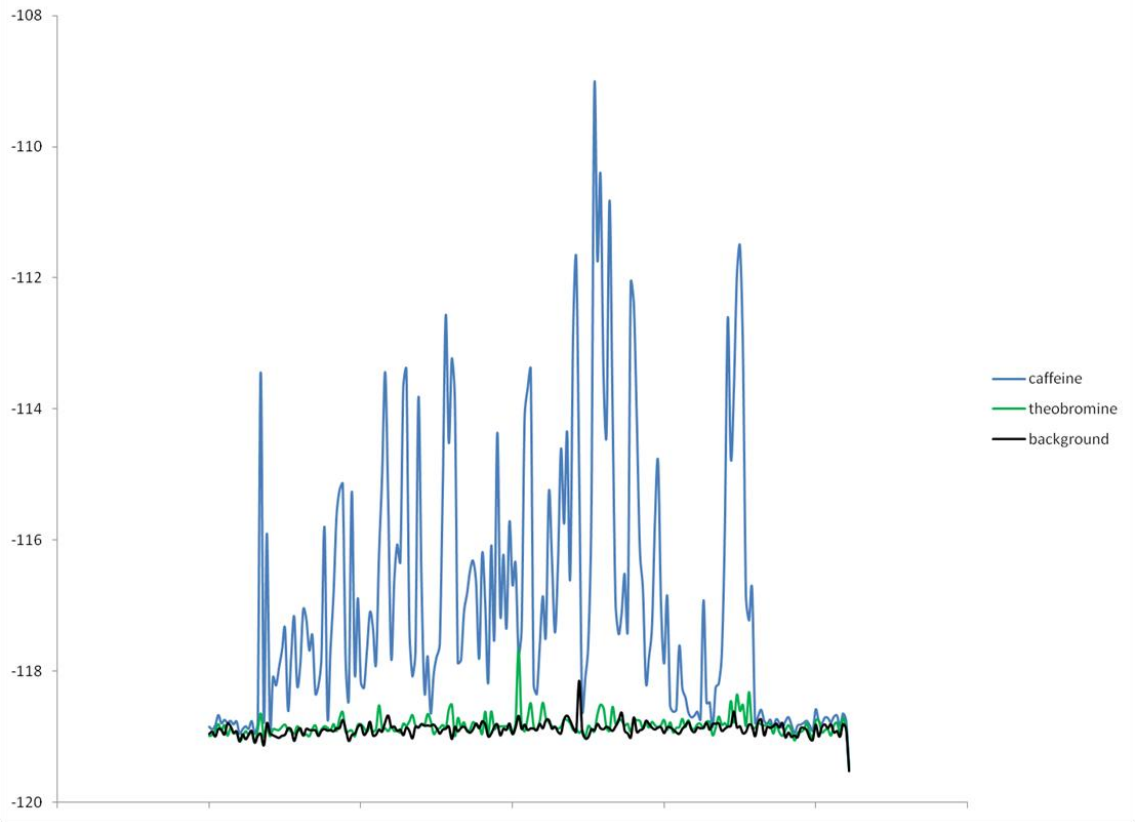
11 Fv 47 F.39 Profile Wall
At ionization wavelength 281.55 nm



Scan along the sample →

11 Fv 47 82-273

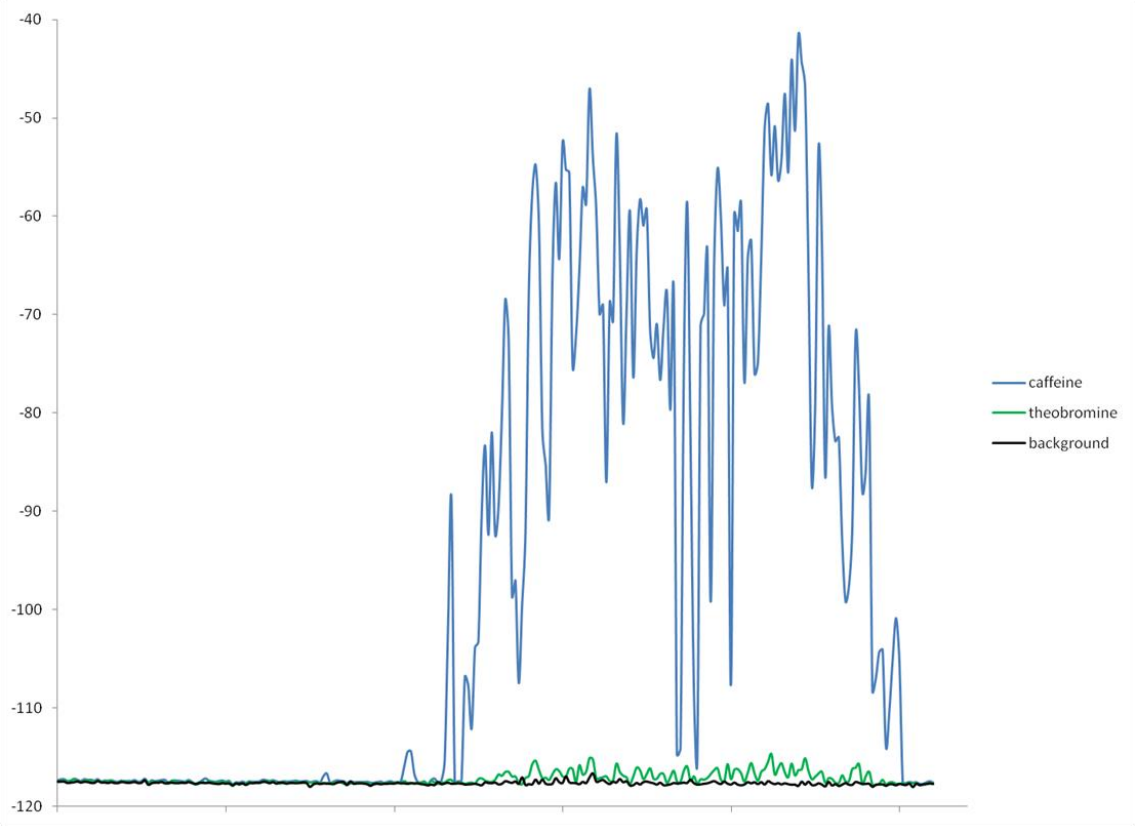
At ionization wavelength 281.55 nm



Scan along the sample →

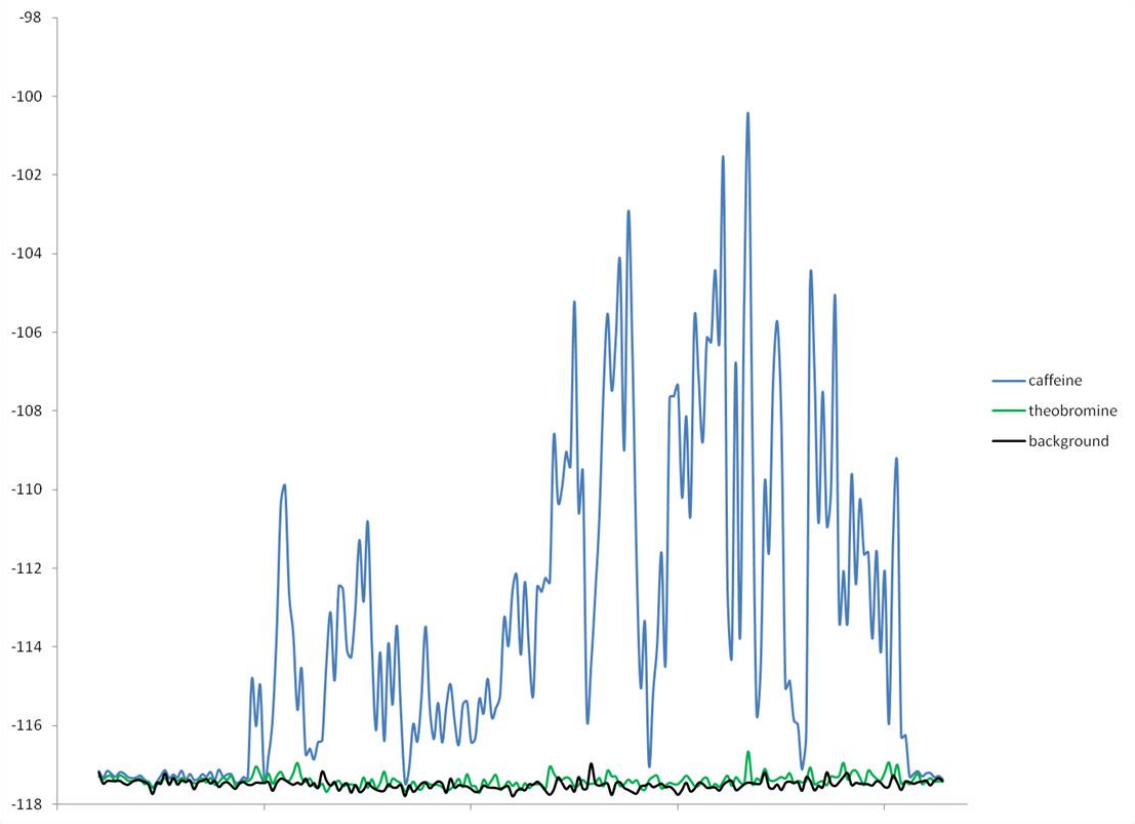
11 Fv 47 F.52.21 ZB NE 1/2

At ionization wavelength 281.55 nm



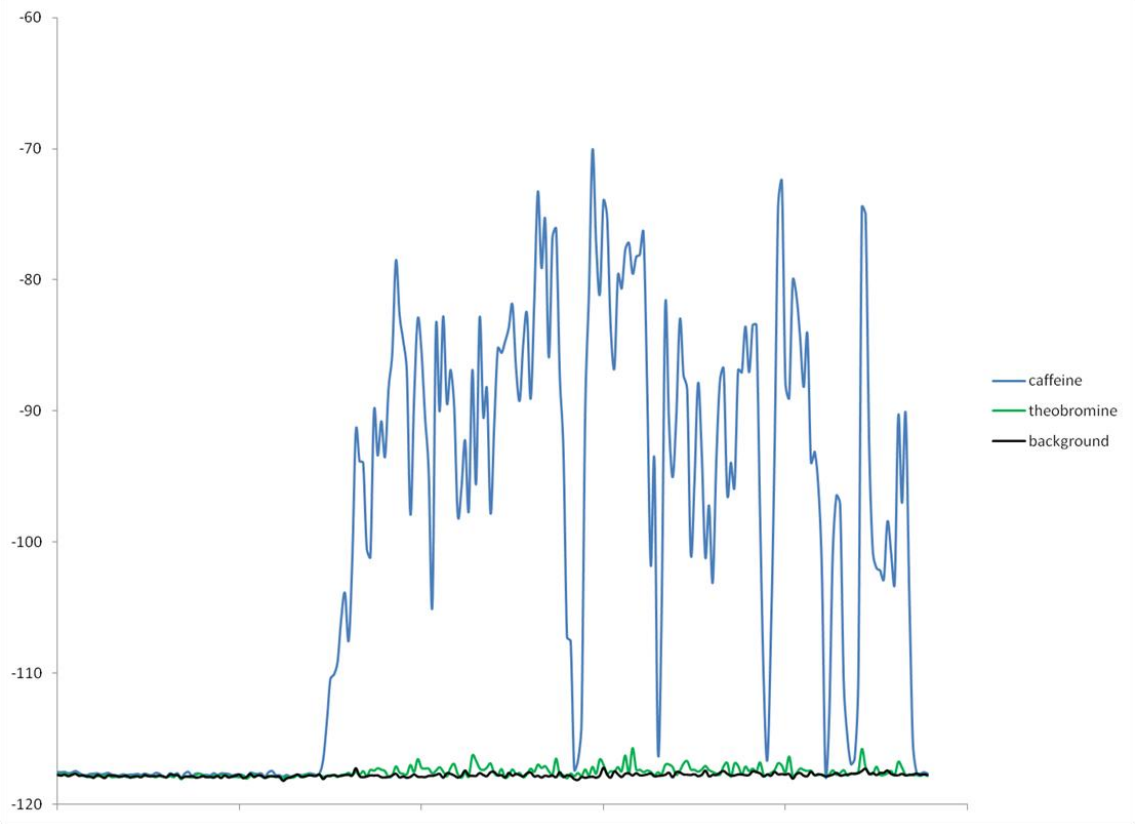
Scan along the sample →

11 Fv-47 lot 31 82A Feature 61
At ionization wavelength 281.55 nm



Scan along the sample →

CW Cooper Feature 12 F.12-8 Zone A SE 1/2
At ionization wavelength 281.55 nm



Scan along the sample →

4. Image of samples prepared for extract and direct desorption analysis